KEK-TH-1449

Stochastic Equations in Black Hole Backgrounds and Non-equilibrium Fluctuation Theorems

Satoshi Iso [1] and Susumu Okazawa [2]

KEK Theory Center, Institute of Particle and Nuclear Studies,
High Energy Accelerator Research Organization(KEK)
and
The Graduate University for Advanced Studies (SOKENDAI),
Oho 1-1, Tsukuba, Ibaraki 305-0801, Japan

Abstract

We apply the non-equilibrium fluctuation theorems developed in the statistical physics to the thermodynamics of black hole horizons. In particular, we consider a scalar field in a black hole background. The system of the scalar field behaves stochastically due to the absorption of energy into the black hole and emission of the Hawking radiation from the black hole horizon. We derive the stochastic equations, i.e. Langevin and Fokker-Planck equations for a scalar field in a black hole background in the $\hbar \to 0$ limit with the Hawking temperature $\hbar\kappa/2\pi$ fixed. We consider two cases, one confined in a box with a black hole at the center and the other in contact with a heat bath with temperature different from the Hawking temperature. In the first case, the system eventually becomes equilibrium with the Hawking temperature while in the second case there is an energy flow between the black hole and the heat bath. Applying the fluctuation theorems to these cases, we derive the generalized second law of black hole thermodynamics. In the present paper, we treat the black hole as a constant background geometry.

[1]E-mail address: satoshi.iso@kek.jp
[2]E-mail address: okazawas@post.kek.jp

Since the paper is also aimed to connect two different areas of physics, non-equilibrium physics and black holes physics, we include pedagogical reviews on the stochastic approaches to the non-equilibrium fluctuation theorems and some basics of black holes physics.

Contents

1 Introduction

The analogy of the space-time with horizons and thermodynamic systems have been extensively investigated, especially, in the black hole thermodynamics [1]. A black hole behaves like a blackbody with the Hawking temperature $T_H = \hbar\kappa/2\pi$ [2], and energy flowing into the black hole can be identified as the entropy increase of the black hole. Here, κ is the surface gravity at the horizon and the entropy of the black hole S_{BH} is proportional to the area of the horizon A as $S_{BH} = A/4G$ in the Einstein-Hilbert theory of gravity. The thermal behavior is essentially quantum mechanical. Furthermore, such a thermal behavior is not restricted to globally-defined horizons like an event horizon of a black hole, but also applicable to local horizons such as the Rindler horizon of a uniformly accelerated observer. At the quantum level, the notion of horizon entropy must have more fundamental meanings, since it gives transition rates of area-changing irreversible processes of black holes, and will be related to the quantum statistical nature of the space-time. Such microscopic views have been proposed in string theory, especially in the approach based on D-brane constructions [3].

Because of the thermodynamic behavior of black holes, a system in a black hole background behaves as a system in contact with a thermal bath. In particular, if we consider a (scalar) field in a black hole background, its effective equation must be described by a stochastic equation with dissipation and quantum noise. The dissipation comes from the classical causal property of the horizon; the black hole horizon absorbs matter and, once they fall in, they cannot come out. The property is the basis of the membrane paradigm of the black hole [4], in which Ohm's law or the Navior Stokes equations hold on the membrane at the (stretched) horizon. On the other hand, the noise term (or fluctuation) comes from the Hawking radiation, which is essentially quantum mechanical and, hence, we need to quantize the system in the black hole background in an appropriate way. The first purpose of the present paper is to derive such a stochastic equation of motion for a scalar field in a black hole background. The stochastic equation of motion of a string is previously derived in [5, 6] based on physical intuition of the Hawking radiation, or in [7] by using an analogy with the Schwinger-Keldysh formalism in the context of AdS/CFT correspondence[8]. Our approach is similar to them, but we obtain the effective equation by explicitly integrating fluctuating degrees of freedom. Namely, we introduce infinitely many variables between the horizon and the stretched horizon and consider them as environmental variables. By integrating them, we can show that the variable at the stretched horizon behaves stochastically with a noise term. Though the environmental variables are living outside of the horizon, they can encode information in the black hole through

choosing the Kruskal vacuum with the regularity condition at the horizon. In this sense, the integration of the environmental variables corresponds to integrating hidden variables in the horizon. The derivation of the Langevin equation is one of our main results.

The second purpose of the paper is to apply the non-equilibrium fluctuation theorem [9]-[10] developed in the statistical physics to the scalar field in the black hole background. In thermodynamic systems, entropy is always increasing (or remaining a constant). But for a mesoscopic system where fluctuations are large, there are nonzero probabilities that the entropy of the system decreases. The fluctuation theorem relates probabilities of entropy decreasing processes to those of entropy increasing ones in terms of the equilibrium thermodynamic quantities. It is a very general theorem that can hold for various dynamical and non-equilibrium systems including classical Hamilton dynamics in contact with a heat bath, stochastic equations with dissipation and noise, or quantum mechanical systems. The Jarzynski equality can be derived from the fluctuation theorem, and the second law of thermodynamics is implied from the Jarzynski equality. We use the word *implied* here because the second law can be derived only if we assume that a system is relaxed to an equilibrium state after a long time. An application of the fluctuation theorem to a scalar field in a black hole background is straightforward once we obtain a stochastic equation of motion. We can derive the generalized second law of black hole thermodynamics, or Green-Kubo formula of the linear response and its nonlinear generalizations.

The paper is organized as follows. In section 2, we briefly review the stochastic approach to thermodynamic systems, Langevin equation and Fokker-Planck equation. An important property of the stochastic equation is that it violates the time reversal symmetry which can be measured by an entropy increase in the path integral. In the next section 3, the fluctuation theorem for a stochastic system is reviewed. It relates the entropy increasing and decreasing probabilities. From the fluctuation theorem, the Jarzynski equality is derived. In section 4, we derive an effective stochastic equation of a scalar field in a black hole background. In deriving the Langevin equation, the quantum property of the vacuum with a regularity condition at the horizon is very important, which is first explained. We then introduce a set of discretized equations of a scalar field near the black hole horizon, and integrate the variables between the horizon and the stretched horizon. The integration leads to an effective stochastic equation for a variable at the stretched horizon. This has the same spirit as deriving a Langevin equation of a system in contact with a thermal bath [11, 12, 13]. In section 5, we apply the fluctuation theorem to the scalar field in a black hole background. We consider two different situations. In the first case, we put the scalar field and the black hole in a box with an insulating wall. By applying the fluctuation theorem, we can derive a relation connecting entropy decreasing

probabilities with increasing ones. The ratio is given by the difference of free energies. From this, the generalized second law of black hole thermodynamics can be derived. In the second case, the wall is assumed to be in contact with a thermal bath of a different temperature which is slightly lower than the Hawking temperature of the black hole. Then there is an energy flow from the black hole to the wall. By applying the fluctuation theorem to it, a linear response theorem of an energy flow to the temperature difference can be obtained. In the appendix A, we review a derivation of the path integral form of the Fokker-Planck equation. In the appendix B, we will discuss the relation between the noise correlation and the flux of the Hawking radiation. In the appendix C, we explain the fluctuation theorem for a steady state and derivations of nonlinear generalizations of Green-Kubo formula.

2 Stochastic Equations of Motion

We first briefly review stochastic approaches to classical statistical systems. In particular, we focus on the path-integral representation (Onsager-Machlup formalism) of the Fokker-Planck equation and emphasize the role of time-reversal symmetry. Readers familiar with non-equilibrium statistical physics can skip this and the next sections.

2.1 The Langevin Equation

The Langevin equation is a phenomenological equation of motion of a particle with a friction term and thermal noise. It is commonly described as

$$m\dot{v} = -\gamma v - \frac{\partial V}{\partial x} + \xi. \tag{2.1}$$

$V(x)$ is an external potential for the particle. γ is the friction coefficient and $\xi(t)$ is a thermal noise (or a random force) which is often assumed to have a Gaussian and white-noise (delta-correlated) distribution

$$\langle \xi(t) \rangle = 0 \ , \quad \langle \xi(t)\xi(t') \rangle = 2\gamma T \delta(t-t'). \tag{2.2}$$

The coefficient $2\gamma T$ is determined to satisfy the equipartition theorem with the temperature T through the fluctuation-dissipation theorem. The noise average $\langle \cdots \rangle$ can be represented by the following path integral

$$\langle F(t) \rangle = \int \mathcal{D}\xi F(t) \exp\left[-\frac{1}{2} \int dt_1 dt_2 \xi(t_1) \frac{\delta(t_1-t_2)}{2\gamma T} \xi(t_2) \right] \tag{2.3}$$

with a normalization condition $\langle 1\rangle = 1$. If necessary, we can easily generalize the noise correlation to an arbitrary colored non-Gaussian noise. An well-known example that can be conveniently described by the Langevin equation is the Brownian motion of a particle or thermal fluctuations of an electric circuit voltage.

2.2 The Fokker-Planck Equation

From the Langevin equation, we can derive another type of a stochastic equation, the Fokker-Planck equation. It describes a dynamical evolution of the probability distribution $P(X,t)$ of observables X at time t. Here X represents the variables $(x, v = \dot{x})$. If the process is Markovian, i.e. the next state is determined only by the present state, the time evolution of P is given by the following Master equation,

$$\partial_t P(X,t|X_0,0) = \int dX' \left[w(X' \to X) P(X',t|X_0,0) - w(X \to X') P(X,t|X_0,0) \right]. \quad (2.4)$$

Here $P(X,t|X_0,0)$ is a conditional probability to find an event $X(t) = X$ that has started from the initial value $X(0) = X_0$ at $t = 0$, i.e. $P(X,t=0|X_0,0) = \delta(X - X_0)$. $w(X' \to X)$ is a transition rate from X' to X, which can be related to the Langevin equation in the following way. The first and the second terms of the right hand side of eq.(2.4) describe an incoming and outgoing fluxes of X respectively.

The Master equation can be brought into the Kramers-Moyal form as

$$\begin{aligned}
&\partial_t P(X,t|X_0,0) \\
&= -\int dr \left[w(X \to X+r) P(X,t|X_0,0) - w(X - r \to X) P(X-r,t|X_0,0) \right] \\
&= -\int dr \left[1 - e^{-r\partial_X} \right] w(X \to X+r) P(X,t|X_0,0) \\
&= \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \partial_X^n \left[C_n(X) P(X,t|X_0,0) \right], \qquad (2.5)
\end{aligned}$$

where we have defined

$$C_n(X) = \int dr r^n w(X \to X+r) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \langle (X(t+\Delta t) - X(t))^n \rangle |_{X(t)=X}. \quad (2.6)$$

In the last line, we have rewritten the n-th moment of the transition rate by a thermal average of an infinitely small variation of the observable X. In this way, we can convert the Langevin equation for dynamical variables to the Fokker-Planck equation for the distribution functions. Here we show an explicit derivation of the Fokker-Planck equation for the simplest Langevin equation (2.1) as a demonstration. Eq.(2.1) can be considered

as a set of first order differential equations for two variables x and $v = \dot{x}$. Then the Kramers-Moyal coefficients up to the second moments are given by

$$\begin{aligned}
C_1(x) &= v \\
C_1(v) &= -\frac{\gamma}{m}v - \frac{1}{m}\frac{\partial V}{\partial x} \\
C_2(x) &= 0 \\
C_2(v) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_t^{t+\Delta t} dt_1 \int_t^{t+\Delta t} dt_2 \langle \dot{v}(t_1)\dot{v}(t_2)\rangle|_{x(t)=x} \\
&= \lim_{\Delta t \to 0} \left(\frac{1}{\Delta t} \int_t^{t+\Delta t} dt_1 \frac{2\gamma T}{m^2} + \mathcal{O}(\Delta t) \right) \\
&= \frac{2\gamma T}{m^2}.
\end{aligned} \tag{2.7}$$

Higher order coefficients vanish in the $\Delta t \to 0$ limit. Now we get the Fokker-Planck equation corresponding to the Langevin equation (2.1);

$$\partial_t P(x, v, t|x_0, v_0, 0) = \partial_x \left(-vP\right) + \partial_v \left[\left(\frac{\gamma}{m}v + \frac{1}{m}\frac{\partial V}{\partial x} \right) P \right] + \partial_v^2 \left(\frac{\gamma T}{m^2} P \right). \tag{2.8}$$

This Fokker-Planck equation has a simple solution

$$P^{\mathrm{st}} \propto e^{-\frac{1}{T}\left(\frac{1}{2}mv^2 + V(x)\right)}. \tag{2.9}$$

Note that both of $-v\partial_x P + \frac{1}{m}\frac{\partial V}{\partial x}\partial_v P$ and $\partial_v \left[\frac{\gamma}{m}vP + \frac{\gamma T}{m^2}\partial_v P\right]$ cancel for P^{st}. It is the well-known Maxwell-Boltzmann distribution for a system in an equilibrium with temperature T, and satisfies the stationarity condition $\partial_t P^{\mathrm{st}} = 0$. The solution satisfies the equilibrium condition, stronger than the stationarity condition.

Here we have used the words "stationary" and "equilibrium" in the following sense. Stationary distributions are solutions to the Fokker-Planck equation satisfying $\partial_t P = 0$. Equilibrium distributions are also stationary but satisfy a stronger condition which is called the detailed balance condition. The most direct definition of the detailed balance condition is given in the language of the Master equation. Due to the definition of stationarity, P^{st} satisfies $\int dX' \left[w(X' \to X)P^{\mathrm{st}}(X') - w(X \to X')P^{\mathrm{st}}(X)\right] = 0$ for arbitrary X. On the other hand, the detailed balance condition is defined as

$$\forall X, X', \quad w(X' \to X)P^{\mathrm{st}}(X') - w(X \to X')P^{\mathrm{st}}(X) = 0. \tag{2.10}$$

To satisfy this condition, the system must have the microscopic time reversal symmetry and can not have a specific arrow of time. In other words, there is no entropy production. In a stationary but non-equilibrium configuration, there is a flow of current in a configuration space (x, v).

The solution of the Fokker-Planck equation can be represented in a path integral form as

$$P(x,t|x_0,0) = \int_{x(0)=x_0}^{x(t)=x} \mathcal{D}x \exp\left[-\tfrac{1}{4\gamma T}\int_0^t dt'\left(m\ddot{x}+\gamma\dot{x}+\tfrac{\partial V}{\partial x}\right)^2\right] \tag{2.11}$$

Its derivation is explained in the appendix A. The "Lagrangian" $L = \frac{1}{4\gamma T}(m\ddot{x}+\gamma\dot{x}+\frac{\partial V}{\partial x})^2$ is called the Onsager-Machlup function [14]. A variation of the Onsager-Machlup function gives the most probable path in the stochastic processes. Apparently, since we have $L \geq 0$, the paths satisfying $L = 0$ are most favored if exist.

The Onsager-Machlup function can be divided into two parts,

$$\frac{1}{4\gamma T}\left(m\ddot{x}+\tfrac{\partial V}{\partial x}\right)^2 + \frac{\gamma}{4T}\dot{x}^2 \tag{2.12}$$

which preserves time reversal symmetry, and a violating term,

$$-\frac{1}{2T}\dot{x}\left(m\ddot{x}+\tfrac{\partial V}{\partial x}\right). \tag{2.13}$$

The latter plays an important role to prove the fluctuation theorem in the next section.

3 Non-equilibrium Identities

The stochastic equations such as the Langevin or the Fokker-Planck equations describe how a system is dynamically relaxed to a stationary or an equilibrium state. Furthermore we can calculate transition amplitudes of a system to one state to another. By using the method reviewed in the previous section, we can calculate a ratio of an entropy decreasing probability to an entropy increasing probability. Since the latter probabilities have always much bigger values, the entropy is always increasing after we take a stochastic average.

In this section we review a derivation of the fluctuation theorem and the Jarzynski equality from the stochastic equations.

3.1 The Fluctuation Theorem

The fluctuation theorem was first discovered in a numerical simulation [9] and gives the ratio of probabilities of an entropy increasing process to that of a decreasing one. The proof of the fluctuation theorem is given for various systems including classical Hamiltonian dynamics [15], stochastic Langevin dynamics [16] and quantum mechanical evolutions [17, 18]. The Jarzynski equality [10] is a relation between non-equilibrium work and

equilibrium free energy difference, and both of them are remarkable discoveries in the recent developments of non-equilibrium statistical physics. In this paper, we concentrate on a system that the evolution is described by a Fokker-Planck equation such as eq.(2.8). The fluctuation theorems can be simply derived and the meaning of entropy production (or a violation of time-reversal symmetry) is clear.

We consider a stochastic system described by the Langevin equation (2.1) or the Fokker-Planck equation (2.8). In order to study a dynamical evolution, we introduce an externally controlled parameter λ_t^F in the potential $V(x;\lambda_t^F)$. By changing the external parameter λ_t^F as a function of t, the corresponding stable state changes accordingly with time. For later convenience, we call the process of changing the external parameter with λ_t^F as the "forward protocol". For example, we may set the minimum position of a harmonic potential as the externally controlled parameter;

$$V(x;\lambda_t^F) = \frac{1}{2}k(x-\lambda_t^F)^2, \tag{3.1}$$

if the position moves linearly in time t, the parameter is given by $\lambda_t^F = v_0 t$. We can also take different protocols e.g. oscillatory or pulse-like etc.

From the path integral representation of the transition rate (2.11), a probability that a sequence of configurations $\Gamma_\tau = \{x(t), t \in [0,\tau] | x(0) = x_{\rm ini}, x(\tau) = x_{\rm fin}\}$ is realized during the time interval $t \in [0,\tau]$ is given by

$$P^F[\Gamma_\tau|x_{\rm ini}] \propto \exp\left[-\tfrac{1}{4\gamma T}\int_{\Gamma_\tau} dt \left(m\ddot{x} + \gamma\dot{x} + \tfrac{\partial V(x;\lambda_t^F)}{\partial x}\right)^2\right]. \tag{3.2}$$

The trajectory Γ_τ represents a sequence of configurations in the forward protocol λ_t^F with the initial configuration $x(0) = x_{\rm ini}$.

We now define a time reversal of the forward protocol λ_t^F, and call it a "reversed protocol" $\lambda_t^R \equiv \lambda_{\tau-t}^F$. We consider a probability $P^R[\Gamma_\tau^*|x_{\rm fin}]$ that the system experiences a reversed trajectory $\Gamma_\tau^* = \{x^*(t) \equiv x(\tau-t), t \in [0,\tau] | x^*(0) = x_{\rm fin}, x^*(\tau) = x_{\rm ini}\}$ in the time-reversed protocol λ_t^R. The reversed trajectory has the initial value $x^*(0) = x_{\rm fin} = x(\tau)$, $\dot{x}^*(0) = -\dot{x}(\tau)$. If the system has time-reversal symmetry, the probability should be the same as the probability $P^F[\Gamma_\tau|x_{\rm ini}]$. But since the stochastic equation violates the symmetry, they will be different. The reversed propability $P^R[\Gamma_\tau^*|x_{\rm fin}]$ is similarly given by

$$\begin{aligned} P^R[\Gamma_\tau^*|x_{\rm fin}] &\propto \exp\left[-\tfrac{1}{4\gamma T}\int_{\Gamma_\tau^*} dt \left(m\ddot{x} + \gamma\dot{x} + \tfrac{\partial V(x;\lambda_t^R)}{\partial x}\right)^2\right] \\ &= \exp\left[-\tfrac{1}{4\gamma T}\int_{\Gamma_\tau} dt' \left(m\ddot{x} - \gamma\dot{x} + \tfrac{\partial V(x;\lambda_{t'}^F)}{\partial x}\right)^2\right]. \end{aligned} \tag{3.3}$$

In the last line, we change a variable from t to $t' = \tau - t$. This change causes a flip of the sign of $\dot{x}$. The ratio of P^F and P^R now becomes

$$\frac{P^F[\Gamma_\tau|x_{\rm ini}]}{P^R[\Gamma^*_\tau|x_{\rm fin}]} = \exp\left[-\tfrac{1}{T}\int_{\Gamma_\tau} dt\dot{x}\left(m\ddot{x} + \tfrac{\partial V(x;\lambda^F_t)}{\partial x}\right)\right]. \tag{3.4}$$

This gives a key property to prove the fluctuation theorem. Time-reversal symmetric terms are canceled between P^F and P^R, and the ratio is given by the entropy production $\dot{S}$ of the stochastic process.

We further need to sum over the initial configurations, $x_{\rm ini}$ and $x_{\rm fin}$ respectively for the forward and the reversed protocols, with appropriate statistical weights. Here we assume that the external parameter is kept fixed at the initial value of each protocol before $t = 0$. Hence the system is in the equilibrium. We therefore multiply P^F or P^R by the Boltzmann weight $P^{\rm eq}(x_{\rm ini})$ or $P^{\rm eq}(x_{\rm fin})$. The ratio of the Boltzmann weights for the initial configurations is given by

$$\begin{aligned}\frac{P^{\rm eq}(x_{\rm ini})}{P^{\rm eq}(x_{\rm fin})} &= \frac{Z(\lambda^F_\tau)}{Z(\lambda^F_0)}\exp\left[-\frac{1}{T}\left(\frac{1}{2}m(\dot{x}^2_{\rm ini} - \dot{x}^2_{\rm fin}) + V(x_{\rm ini};\lambda^F_0) - V(x_{\rm fin};\lambda^F_\tau)\right)\right] \\ &= \exp\left[\frac{1}{T}\int_{\Gamma_\tau} dt\left(m\dot{x}\ddot{x} + \dot{x}\frac{\partial V(x;\lambda^F_t)}{\partial x} + \dot{\lambda}^F_t\frac{\partial V(x;\lambda^F_t)}{\partial \lambda^F_t}\right) - \frac{\Delta F}{T}\right],\end{aligned} \tag{3.5}$$

where ΔF is a difference of the free energies $F(\lambda) = -T\log Z(\lambda)$ of equilibrium states at $\lambda = \lambda^F_0$ and $\lambda = \lambda^F_\tau$,

$$\Delta F = F(\lambda^F_\tau) - F(\lambda^F_0). \tag{3.6}$$

Combining the two ratios eq.(3.4) and eq.(3.5), we get the following relation,

$$\frac{P^F[\Gamma_\tau|x_{\rm ini}]P^{\rm eq}(x_{\rm ini})}{P^R[\Gamma^*_\tau|x_{\rm fin}]P^{\rm eq}(x_{\rm fin})} = \exp\left(R[\Gamma_\tau]\right). \tag{3.7}$$

Here we have defined $R[\Gamma_\tau]$ and $W[\Gamma_\tau]$ as

$$R[\Gamma_\tau] \equiv \frac{1}{T}\int_{\Gamma_\tau} dt\dot{\lambda}^F_t\frac{\partial V(x;\lambda^F_t)}{\partial \lambda^F_t} - \frac{\Delta F}{T} \equiv W[\Gamma_\tau] - \frac{\Delta F}{T} \tag{3.8}$$

which measures the entropy production in the trajectory Γ_τ and the work exerted on the system.

As a simple example, for the potential $V(x;\lambda^F_t) = k(x - v_0 t)^2/2$, we have

$$R[\Gamma_\tau] = -\frac{1}{T}\int_{\Gamma_\tau} dt v_0 k(x(t) - v_0 t). \tag{3.9}$$

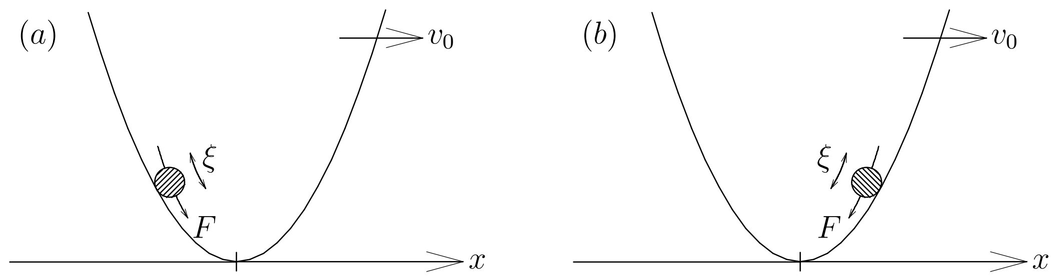

Figure 1: (a) A schematic illustration of motion of a particle in a potential $V(x;\lambda_t^F) = \frac{1}{2}k(x - v_0 t)^2$. This picture shows a *natural* configuration with $(x(t) - v_0 t) < 0$. It gives a positive value of $R[\Gamma_\tau]$. (b) A noise ξ rarely pushes a particle to the opposite side beyond the minimum point $x(t) = v_0 t$. Since $(x(t) - v_0 t) > 0$, it gives a negative value of $R[\Gamma_\tau]$

The term, velocity times force, gives a work exerted on the system. If we neglected the fluctuation of the particle, $x(t) - v_0 t$ would always have a negative sign, and $R[\Gamma_\tau]$ would always increase. It is consistent with a naive picture. However in a mesoscopic system, fluctuations can grow larger and $x(t) - v_0 t$ can have a positive sign. Then the particle overshoots the equilibrium point $\partial_x V = 0$ to the positive side and $R[\Gamma_\tau]$ becomes negative. Such a negative value of $R[\Gamma_\tau]$ indicates that the system exerts work onto outside and it gives a negative entropy production.

From the equation (3.7), by integrating all the paths of the configurations, we can derive the fluctuation theorem in the final form as

$$\begin{aligned}
\rho^F(R_\tau) &\equiv \int \mathcal{D}x P^F[\Gamma_\tau|x_{\rm ini}]P^{\rm eq}(x_{\rm ini})\delta(R_\tau - R[\Gamma_\tau]) \\
&= \int \mathcal{D}x P^R[\Gamma_\tau^*|x_{\rm fin}]P^{\rm eq}(x_{\rm fin})e^{R[\Gamma_\tau]}\delta(R_\tau - R[\Gamma_\tau]) \\
&= e^{R_\tau} \int \mathcal{D}x P^R[\Gamma_\tau^*|x_{\rm fin}]P^{\rm eq}(x_{\rm fin})\delta(R_\tau + R[\Gamma_\tau^*]) \\
&= e^{R_\tau}\rho^R(-R_\tau). \qquad (3.10)
\end{aligned}$$

The first line is the definition of $\rho^F(R_\tau)$, i.e. the probability to get the entropy production R_τ within the interval $[0,\tau]$. We use the relation (3.7) in the second line. In the third equality the relation $R[\Gamma_\tau^*] = -R[\Gamma_\tau]$ is used. Since the quantity R_τ measures the entropy production in the interval, we see that entropy decreasing probabilities are related to increasing ones. They are exponentially suppressed, but exist with nonzero probabilities.

3.2 The Jarzynski Equality

By integrating the fluctuation theorem over the entropy production, we can construct an equality, so called the Jarzynski equality [19].

$$\begin{aligned}\int_{-\infty}^{\infty} dR_\tau \rho^F(R_\tau) e^{-R_\tau} &= \int_{-\infty}^{\infty} dR_\tau \rho^R(-R_\tau) \\ \Rightarrow \langle e^{-R_\tau} \rangle &= 1. \end{aligned} \tag{3.11}$$

We have defined the average as

$$\langle F(R_\tau) \rangle = \int_{-\infty}^{\infty} dR_\tau \rho^F(R_\tau) F(R_\tau) = \int \mathcal{D}x P^F[\Gamma_\tau | x_{\text{ini}}] P^{\text{eq}}(x_{\text{ini}}) F(R[\Gamma_\tau]). \tag{3.12}$$

The Jarzynski equality (3.11) states that the weighted sum of e^{-R_τ} over all possible non-equilibrium processes with an externally controlled potential gives an unity. In terms of the work exerted on the system $W[\Gamma_\tau]$ and the free energy difference, we can relate an average work done in non-equilibrium processes to the equilibrium free energy difference [10] as

$$\langle e^{-\frac{W}{T}} \rangle = e^{-\frac{\Delta F}{T}}. \tag{3.13}$$

From this, by using the Jensen inequality $\langle e^x \rangle \geq e^{\langle x \rangle}$, we get an inequality;

$$\langle W \rangle - \Delta F \geq 0. \tag{3.14}$$

This indicates the second law of thermodynamics. The Jarzynski equality simply states that there must exist microscopic processes with a large negative entropy production to satisfy the equality, and the probability is characterized by the equilibrium quantity of the free energy difference.

Some comments are in order. First the notion of entropy is usually defined for a thermal system after taking an average. So it may be appropriate to use a word, an entropy function, instead of the entropy for each microscopic configuration. The second comment is that in the above *derivation* of the second law we have implicitly in mind that the above free energy difference is the difference between the initial and the final free energies. It is justified if the system is relaxed to an equilibrium state with the external parameter at $t = \tau$ after a long time interval. Since the system is in contact with a large heat bath with temperature T, the relaxed state coincides with the equilibrium state at the temperature. If this is the case, the second law of thermodynamics is derived from the Jarzynski equality. In the present proof of the fluctuation theorem, we have used the stochastic approach

and the system explicitly violates the time-reversal symmetry. Then such a relaxation can occur. But if we starts from the original unitary quantum mechanical evolution, the system cannot be thermalized in an exact sense. In applying the fluctuation theorem to the information paradox of a black hole, such considerations are inevitable.

An alternative expression of the fluctuation theorem is obtained by using a generating function. We define the generating function for R_τ as

$$Z^F(\alpha_\tau) = \ln\left(\int_{-\infty}^{\infty} dR_\tau e^{i\alpha_\tau R_\tau}\rho^F(R_\tau)\right). \tag{3.15}$$

Derivatives of $Z^F(\alpha_\tau)$ give connected correlators of the entropy production R_τ in a situation of the forwardly varying parameter. One easily gets the following relation between $Z^F(\alpha_\tau)$ and $Z^R(\alpha_\tau)$ from the fluctuation theorem as

$$\begin{aligned} Z^F(\alpha_\tau) &= \ln\left(\int_{-\infty}^{\infty} dR_\tau e^{i\alpha_\tau R_\tau} e^{R_\tau}\rho^R(-R_\tau)\right) \\ &= \ln\left(\int_{-\infty}^{\infty} dx e^{ix(i-\alpha_\tau)}\rho^R(x)\right) \\ &= Z^R(i-\alpha_\tau). \end{aligned} \tag{3.16}$$

We have used the equation (3.10) in the first line. In the second line, we changed a variable R_τ to $x = -R_\tau$. If the forward and the reversed protocols are identical i.e. $\lambda_t^F = \lambda_{\tau-t}^F$, we get a simpler relation $Z(\alpha_\tau) = Z(i-\alpha_\tau)$.

Finally, we give a comment on our assumption of the initial distribution. We have assumed that the initial distribution is an equilibrium one. This condition can be easily relaxed to a steady state. More generally, if the initial distributions for x_{ini} and x_{fin} are $P^{\text{st}}(x_{\text{ini}})$ and $P^{\text{st}}(x_{\text{fin}})$ respectively, we can define an entropy production as

$$R[\Gamma_\tau] \equiv \ln\left(\frac{P^F[\Gamma_\tau|x_{\text{ini}}]P^{\text{st}}(x_{\text{ini}})}{P^R[\Gamma_\tau^*|x_{\text{fin}}]P^{\text{st}}(x_{\text{fin}})}\right). \tag{3.17}$$

Then we get the fluctuation theorem in the form; $\rho^F(R_\tau)/\rho^R(-R_\tau) = e^{R_\tau}$. The choice of initial distributions is arbitrary, but the problem is that we usually do not know an explicit form of the distribution function of a steady state P^{st}. The steady state fluctuation theorems are reviewed in the appendix C. By using it, we can derive the Green-Kubo formula and its non-linear generalizations.

4 Langevin equation in a Black Hole Background

In this section we derive a stochastic equation for a scalar field in the black hole background. We take $\hbar \to 0$ limit with the Hawking temperature $\hbar\kappa/2\pi$ fixed. Since the energy is absorbed into the black hole, a dissipation term is induced at the horizon. The classical equation is furthermore modified by the quantum effect, i.e. the Hawking radiation from the black hole. Because of these effects, the equation of motion in the black hole background is described by a stochastic Langevin equation with a quantum noise and a classical dissipation terms. We first review the basics of black holes and the Hawking radiation, and then derive the Langevin equation of a scalar field in the black hole background.

4.1 Space-time Structure of Black Holes

First we summarize some basic facts of the space-time structure of black holes. (For a review, see for example [20].) Here, we consider a spherically symmetric neutral black hole, the Schwarzschild black hole. It is a solution to the Einstein equation in vacuum with a zero cosmological constant and the metric is given by

$$\begin{aligned} ds^2 &= -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2 d\Omega^2, \\ f(r) &= 1 - \frac{2GM}{r}, \quad d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2. \end{aligned} \tag{4.1}$$

M is the mass of the black hole and the only parameter of the solution. The solution is asymptotically flat; it approaches the flat metric at the space-like infinity $r \to \infty$. It has time-translation symmetry and the associated time-like Killing vector is given by $\xi = \partial_t$. A Killing horizon is defined as a null hypersurface on which there is a null Killing vector. In the present case, it is given by the condition $g(\xi, \xi) = -f(r) = 0 \leftrightarrow r = r_H = 2GM$. The surface gravity κ is defined on the Killing horizon via the relation

$$\nabla_\xi \xi = \kappa \xi. \tag{4.2}$$

A direct calculation shows that $\kappa = f'(r)/2|_{r=r_H} = 1/4GM$ for the Schwarzschild black hole.

There are several different definitions of horizons. An apparent horizon is a more general concept and defined locally as the most outer trapped null surface. It does not need a time-like Killing vector as the Killing horizon, but it is defined in an observer-dependent way. An event horizon is defined in a global way as a boundary of the past

light cone of the future infinity. Mathematically a black hole is defined as a set that is not contained in the past light cone of the future infinity. For the Schwarzschild black hole, all the definitions of the horizon coincide, but they are different for dynamical black holes. In applying non-equilibrium statistical physics to the dynamics of black holes, we need to pay special attentions to the differences. In the present paper, however, since we consider an eternal black hole as a background space-time, their differences are not essential.

The coordinates used in eq.(4.1) is called the Schwarzschild coordinates. The singularity of the metric at the horizon $r = r_H$ is not physical, and can be removed by using other coordinates, such as the Kruskal (-Szekeres) coordinates (U, V)

$$U = -\frac{1}{\kappa}e^{-\kappa(t-r_*)}, \quad V = \frac{1}{\kappa}e^{\kappa(t+r_*)} \tag{4.3}$$

$$r_* \equiv \int \frac{dr}{f(r)} = r + r_H \log(\frac{r}{r_H} - 1). \tag{4.4}$$

Here r_* is the tortoise coordinate and takes $-\infty < r_* < \infty$ between the horizon and the spacial infinity. In terms of the Kruskal coordinates, the metric of the Schwarzschild black hole becomes regular at the horizon;

$$ds^2 = -\frac{r_H}{r}e^{-\frac{r}{r_H}}dUdV + r^2 d\Omega^2. \tag{4.5}$$

At the price of removing the coordinate singularity, the asymptotically flatness is unclear in these coordinates. We will impose regularity conditions on physical quantities at the horizon in the Kruskal coordinates.

Figure 2 is the Penrose diagram of the Schwarzschild black hole, which captures the causal structure of the space-time. The vertical and horizontal axises correspond to the Kruskal time $T = (V + U)/2$, and the Kurskal radius $R = (V - U)/2$. In contrast to the Schwarzschild coordinates, the Kruskal coordinates are regular beyond the horizon ($r = r_H$), and can be extended to the maximally extended Schwarzschild space-time ($-\infty < U, V < \infty$). The original Schwarzschild coordinates ($-\infty < t < \infty, r_H < r < \infty$), on the contrary, can cover only the region I in fig.2. We define (t, r_*) coordinates in other regions. For example, in the region II, we can define them by the relations $U = e^{-\kappa(t-r_*)}/\kappa, V = -e^{\kappa(t+r_*)}/\kappa$. In the Kruskal coordinates, the space-time is separated by the future and past event horizons ($U = 0$ and $V = 0$ respectively) into four regions. There are four possible combinations of signature of U and V as shown in the table 4.1.

Finally we note that the time-like Killing vector $\xi = \partial_t$ is written as $\xi = \kappa(V\partial_V - U\partial_U) = \kappa R\partial_T$ in the Kruskal coordinates and, therefore, the directions of time are opposite in the region I and II. We have drawn the directions of ξ in fig.2.

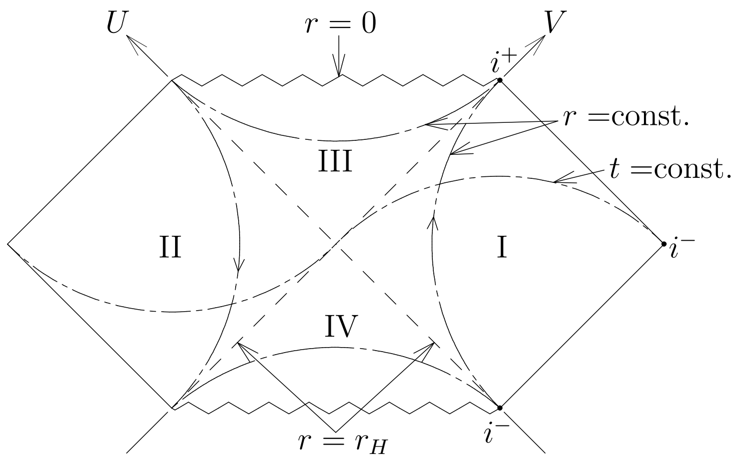

Figure 2: A point in the diagram represents a two dimensional sphere with radius r at time t. r-constant and t-constant surfaces are depicted. Arrows on the r-constant surfaces indicate the flow of the time-like Killing vector. They have opposite directions in the region I and Ⅱ. The singularity at $r = 0$ is drawn by zigzag lines in the diagram. Event horizons are located at $r = r_H$ and separate the space-time into four distinct regions. i^+, i^- and i_0 are the future, past and spatial infinities.

I	$U < 0, V > 0$	$r > r_H$
Ⅱ	$U > 0, V < 0$	$r > r_H$
Ⅲ	$U > 0, V > 0$	$r < r_H$
Ⅳ	$U < 0, V < 0$	$r < r_H$

Table 1: Four regions of maximally extended Schwarzschild space-time

4.2 Field Theory in the Black Hole Background and the Hawking Radiation

We briefly summarize the quantum field theories in the black hole background. For a comprehensive review, see e.g. [21]. The action of a massive scalar field in the maximally extended Schwarzschild space-time is given by a sum of the fields in the right wedge (region I) and in the left wedge (region Ⅲ). In each region, the action is given by

$$S = \int d^4x\sqrt{-g}\frac{1}{2}\left(g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + m^2\phi^2\right) = \sum_{l,m}\int dtdr_*\phi_{(l,m)}\left[\partial_t^2 - \partial_{r_*}^2 + V_l(r)\right]\phi_{(l,m)}. \tag{4.6}$$

where we have decomposed the field into partial waves

$$\phi(t,r,\Omega) = \sum_{l,m} \frac{\phi_{(l,m)}(t,r)}{r} Y_{l,m}(\Omega), \tag{4.7}$$

and defined the effective potential for each partial wave with an angular momentum l,

$$V_l(r) = f(r)\left(\frac{l(l+1)}{r^2} + \frac{\partial_r f(r)}{r} + m^2\right). \tag{4.8}$$

The equation of motion of the scalar field is given by

$$\left[\partial_t^2 - \partial_{r_*}^2 + V_l(r)\right]\phi_{R,L(l,m)} = 0. \tag{4.9}$$

Both in the asymptotically flat region ($r \to \infty$) and in the near horizon region ($r \to r_H$), the potential V_l vanishes and the equation of motion is reduced to the free field equation. Thus, in the near horizon region, the classical solutions are approximately given by

$$u_k^R(t,r) = \begin{cases} \frac{1}{\sqrt{4\pi\omega_k}} e^{-i\omega_k t + ikr_*} & \text{(in R)} \\ 0 & \text{(in L)} \end{cases} \tag{4.10}$$

$$u_k^L(t,r) = \begin{cases} 0 & \text{(in R)} \\ \frac{1}{\sqrt{4\pi\omega_k}} e^{i\omega_k t + ikr_*} & \text{(in L).} \end{cases} \tag{4.11}$$

and their complex conjugates. Here $\omega_k = +|k| > 0$. The sign difference in front of $i\omega_k t$ in R, L follows the convention of [21]. With this convention, these fields are positive frequency modes with respect to the time-like Killing vector, ∂_t in R and $-\partial_t$ in L, satisfying $\mathcal{L}_{\pm\partial_t} u_k = -i\omega_k u_k$. The complex conjugates $(u_k^{R,L})^*$ are the negative frequency modes (in the above sense) satisfying $\mathcal{L}_{\pm\partial_t} u_k^* = +i\omega_k u_k^*$. They are orthonormal with respect to the following Klein-Gordon inner product,

$$\begin{aligned}(f,g) &\equiv i\int_{\Sigma_t} d^3x \sqrt{h_{\Sigma_t}}\,(f^*\partial_t g - \partial_t f^* g) \\ &= i\sum_{l,m}\int dr_*\left(f^*_{(l,m)}\partial_t g_{(l,m)} - \partial_t f^*_{(l,m)} g_{(l,m)}\right).\end{aligned} \tag{4.12}$$

The integration is performed on a constant time slice Σ_t, but it can be generalized to any space-like surface Σ and the choice of the integration surface does not change the value of the inner product. The field $\phi_{(l,m)}$ can be expanded in terms of the classical solutions in the Schwarzschild coordinates in the near horizon region as follows;

$$\phi_{(l,m)} = \int \frac{dk}{\sqrt{4\pi\omega_k}}\left[a^R_{k(l,m)} u^R_k + (a^R_{k(l,m)})^\dagger (u^R_k)^* + a^L_{k(l,m)} u^L_k + (a^L_{k(l,m)})^\dagger (u^L_k)^*\right]. \tag{4.13}$$

We will omit the suffixes (l,m) of creation and annihilation operators for simplicity in the followings.

In the Kruskal coordinates near the horizon, the equation of motion becomes $\partial_U\partial_V\phi_{(l,m)} = (\partial_T^2 - \partial_R^2)\phi_{(l,m)} = 0$. So we may define another basis of functions

$$u_p^K(T,R) = \frac{1}{\sqrt{4\pi E_p}} e^{-iE_pT+ipR}, \tag{4.14}$$

where $E_p = +|p| > 0$. They are positive frequency modes with respect to the Kruskal time. In terms of them, the field can be expanded as

$$\phi_{(l,m)} = \int \frac{dp}{\sqrt{4\pi E_p}} \left[b_p u_p^K + (b_p)^\dagger (u_p^K)^* \right]. \tag{4.15}$$

In contrast to the wave functions (4.11), they are defined globally in the whole space-time.

In order to relate two different definitions of the creation and annihilation operators in the Kruskal and Schwarzschild coordinates and to express the Kruskal vacuum $b_k|0\rangle_K = 0$ as a Fock state constructed on the Schwarzschild vacuum $a_k^{R,L}|0\rangle_{R,L} = 0$, we look at the analyticity properties of the functions [22]. The positive frequency wave function u_p^K in the Kruskal coordinates with $p > 0$ (or $p < 0$) is an analytic function in the lower half U (or V) plane since $u_p^K \sim e^{-iE_pU}$ (or $u_p^K \sim e^{-iE_pV}$). On the other hand, since $e^{ikr_*} = (r/r_H - 1)^{ik} e^{ikr}$, there is a phase jump when it crosses the horizon. So we need to combine the positive and negative frequency wave functions in the Schwarzschild coordinates to construct a wave function with the same analyticity property as u_p^K. They were obtained by Unruh [22] as

$$\begin{cases} u_k^{(1)} = \frac{1}{\sqrt{2\sinh\frac{\pi\omega_k}{\kappa\hbar}}} \left(e^{\frac{\pi\omega_k}{2\kappa\hbar}} u_k^R + e^{-\frac{\pi\omega_k}{2\kappa\hbar}} (u_{-k}^L)^* \right) \\ u_k^{(2)} = \frac{1}{\sqrt{2\sinh\frac{\pi\omega_k}{\kappa\hbar}}} \left(e^{-\frac{\pi\omega_k}{2\kappa\hbar}} (u_{-k}^R)^* + e^{\frac{\pi\omega_k}{2\kappa\hbar}} u_k^L \right). \end{cases} \tag{4.16}$$

These combinations are analytic in the lower half plane of U or V. In the following we set $\hbar = 1$ for notational simplicity. Such analyticity property can be easily checked. For example, $u_k^{(1)}$ with a positive k can be rewritten as an analytic function of U

$$\begin{aligned} u_k^{(1)} &\propto u_k^R + e^{-\frac{\pi\omega_k}{\kappa}} (u_{-k}^L)^* \\ &\propto (-\kappa U)^{\frac{i\omega_k}{\kappa}}, \end{aligned} \tag{4.17}$$

if it is analytically continued from the region I of the right wedge ($U < 0$) to the region II of the left wedge ($U > 0$) through the lower half of the U plane by the transformation $U \to Ue^{i\pi}$. Hence the combination is analytic in the lower half plane of U. For $k < 0$, $u_k^{(1)} \propto$

$(\kappa V)^{-\frac{i\omega_k}{\kappa}}$ and it is analytic in the lower half plane of V as e^{-iE_pV}. In the classical limit where $\hbar \to 0$, $u_k^{(1)}$ becomes a positive frequency mode in the Schwarzschild coordinates $e^{-i\omega_k(t\mp r_*)}$ and localized in the region I. Similarly, $u_k^{(2)}$ with a positive momentum $k>0$ is written as an analytic function of the lower half plane of V, $(\kappa V)^{i\omega_k/\kappa}$ while, for a negative k, it is analytic in the lower half plane of U and written as $(-\kappa U)^{-i\omega_k/\kappa}$. It behaves as a negative frequency mode in the Schwarzshchild coordinates but localized mostly in the left wedge in the classical limit. They penetrate into the right wedge by quantum effects. In this sense, $u_k^{(1)}$ is *classical* while $u_k^{(2)}$ is *quantum* in the right wedge.

The scalar field can be expanded in terms of these modes as

$$\phi_{(l,m)} = \int \frac{dk}{\sqrt{4\pi\omega_k}} \left[c_k^{(1)} u_k^{(1)} + (c_k^{(1)})^\dagger (u_k^{(1)})^* + c_k^{(2)} u_k^{(2)} + (c_k^{(2)})^\dagger (u_k^{(2)})^* \right]. \tag{4.18}$$

The Kruskal vacuum ($b_p|0\rangle_K = 0$) is equivalently given by the conditions, $c_k^{(1)}|0\rangle_K = c_k^{(2)}|0\rangle_K = 0$. The annihilation operators in the Schwarzschild coordinates a_k^R and a_k^L can be expressed as a linear combination of $c_k^{(1)}$ and $c_k^{(2)}$ as

$$\begin{cases} a_k^R = \frac{1}{\sqrt{2\sinh\frac{\pi\omega_k}{\kappa}}}\left(e^{\frac{\pi\omega_k}{2\kappa}} c_k^{(1)} + e^{-\frac{\pi\omega_k}{2\kappa}} (c_{-k}^{(2)})^\dagger\right) & = \sqrt{1+n(\omega_k)}\,c_k^{(1)} + \sqrt{n(\omega_k)}\,(c_{-k}^{(2)})^\dagger \\ a_k^L = \frac{1}{\sqrt{2\sinh\frac{\pi\omega_k}{\kappa}}}\left(e^{\frac{\pi\omega_k}{2\kappa}} c_k^{(2)} + e^{-\frac{\pi\omega_k}{2\kappa}} (c_{-k}^{(1)})^\dagger\right) & = \sqrt{1+n(\omega_k)}\,c_k^{(2)} + \sqrt{n(\omega_k)}\,(c_{-k}^{(1)})^\dagger. \end{cases} \tag{4.19}$$

where $n(\omega_k) = 1/(e^{2\pi\omega_k/\kappa}-1)$. Hence the Kruskal and the Schwarzschild operators are related by the Bogoliubov transformation,

$$\begin{pmatrix} a_k^R \\ (a_{-k}^L)^\dagger \end{pmatrix} = \begin{pmatrix} \sqrt{1+n(\omega_k)} & \sqrt{n(\omega_k)} \\ \sqrt{n(\omega_k)} & \sqrt{1+n(\omega_k)} \end{pmatrix} \begin{pmatrix} c_k^{(1)} \\ (c_{-k}^{(2)})^\dagger \end{pmatrix} \equiv U_k \begin{pmatrix} c_k^{(1)} \\ (c_{-k}^{(2)})^\dagger \end{pmatrix}. \tag{4.20}$$

The transformation can also be represented as

$$\begin{aligned} c_k^{(1)} &= e^{-iG} a_k^R e^{iG}, \quad c_{-k}^{(2)} = e^{-iG} a_{-k}^L e^{iG}, \\ G &= i\int \frac{dk}{(2\pi)2\omega_k}\theta_k \left((a_k^R)^\dagger (a_{-k}^L)^\dagger - a_k^R a_{-k}^L\right), \\ \sinh^2\theta_k &\equiv n(\omega_k). \end{aligned} \tag{4.21}$$

From this transformation law, we can read off the relation between Kruskal vacuum and Schwarzschild vacuum as

$$|0\rangle_K = e^{-iG}|0\rangle_R|0\rangle_L \tag{4.22}$$

$$= \prod_k \frac{1}{\cosh\theta_k} \sum_{n=0}^{\infty} e^{-\frac{\beta\omega_k}{2}n_k} |n_k^R\rangle|n_{-k}^L\rangle. \tag{4.23}$$

Note that the Fock space in the left wedge $|n_k^L\rangle$ is constructed on a Minkowski vacuum with the backward time direction $(-t)$.

The expectation value of the Schwarzschild number operators $(a_k^R)^\dagger a_k^R$ in the Kruskal vacuum $|0\rangle_K$ is given by

$$\begin{aligned}
{}_K\langle 0|(a_k^R)^\dagger a_k^R|0\rangle_K &= \frac{1}{2\sinh\frac{\pi\omega_k}{\kappa}} e^{-\frac{\pi\omega_k}{\kappa}} \; {}_K\langle 0|c_{-k}^{(2)}(c_{-k}^{(2)})^\dagger|0\rangle_K \\
&= \frac{1}{e^{\frac{2\pi\omega_k}{\kappa}}-1} = n(\omega_k).
\end{aligned} \tag{4.24}$$

This is the thermal distribution of the Hawking radiation [2], and characterized by the temperature $T_H = \kappa\hbar/2\pi$. Note that the thermal spectrum in the right wedge is created by the effect of the field $u_k^{(2)}$, which is classically localized in the left wedge but penetrates into the right quantum mechanically.

For a generic operator $\hat{\mathcal{O}}_R = \hat{\mathcal{O}}_R(a^R, (a^R)^\dagger)$ which is made of only a^R and $(a^R)^\dagger$, its expectation value ${}_K\langle 0|\hat{\mathcal{O}}_R|0\rangle_K$ can be interpreted as a thermal average. Such a thermal behavior can be generalized to products of operators, such as ${}_K\langle 0|\hat{\mathcal{O}}_L\hat{\mathcal{O}}_R|0\rangle_K$, made of both the right and left creation (annihilation) operators. Its expectation value can be interpreted as a Schwinger-Keldysh correlator.

First let us consider ${}_K\langle 0|\hat{\mathcal{O}}_R|0\rangle_K$. Since the Kruskal vacuum is represented as (4.21), one has

$$\begin{aligned}
{}_K\langle 0|\hat{\mathcal{O}}_R|0\rangle_K &= \prod_k \frac{1}{\cosh^2\theta_k}\sum_{n=0}^{\infty}\langle n_k^R|e^{-\beta\omega_k n_k}\hat{\mathcal{O}}_R|n_k^R\rangle \\
&= \mathrm{Tr}_R\left[\frac{e^{-\beta H_R}}{Z_R}\hat{\mathcal{O}}_R\right].
\end{aligned} \tag{4.25}$$

Here, the Hamiltonian and the partition function are defined by

$$\begin{aligned}
H_R &= \int \frac{dk}{2\pi}\omega_k (a_k^R)^\dagger a_k^R, \\
Z_R &= \mathrm{Tr}\left[e^{-\beta H_R}\right] = \prod_k \sum_{n=0}^{\infty} e^{-\beta\omega_k n_k^R} = \prod_k \cosh^2\theta_k.
\end{aligned} \tag{4.26}$$

Hence ${}_K\langle 0|\hat{\mathcal{O}}_R|0\rangle_K$ can be interpreted as a thermal average of the operator $\hat{\mathcal{O}}_R$ at the Hawking temperature T_H.

For a product of left and right operators, the expectation value in the Kruskal vacuum

is given by

$$
\begin{aligned}
{}_K\langle 0|\hat{\mathcal{O}}_L\hat{\mathcal{O}}_R|0\rangle_K &= \prod_{k,k'}\frac{1}{\cosh\theta_k\cosh\theta_{k'}}\sum_{m,n=0}^{\infty} e^{-\frac{\beta}{2}(\omega_k m_k+\omega_{k'}n_{k'})}\langle m_k^R|\langle m_{-k}^L|\hat{\mathcal{O}}_L\hat{\mathcal{O}}_R|n_{k'}^R\rangle|n_{-k'}^L\rangle \\
&= \prod_{k,k'}\frac{1}{\cosh\theta_k\cosh\theta_{k'}}\sum_{m,n=0}^{\infty} e^{-\frac{\beta}{2}(\omega_k m_k+\omega_{k'}n_{k'})}\langle m_k^R|\hat{\mathcal{O}}_R|n_{k'}^R\rangle\langle m_{-k}^L|\hat{\mathcal{O}}_L|n_{-k'}^L\rangle.
\end{aligned}
\tag{4.27}
$$

In order to express it as an expectation value in the right wedge Fock space, we first rewrite the expectation value $\langle n_{-k}^L|\hat{\mathcal{O}}_L|m_{-k'}^L\rangle$ in terms of the operator in the right wedge as follows,

$$
\langle m_{-k}^L|\hat{\mathcal{O}}_L|n_{-k'}^L\rangle = \langle n_{k'}^R|\hat{\mathcal{O}}_L^{\vee}|m_k^R\rangle, \tag{4.28}
$$

Here we have defined the operator $\hat{\mathcal{O}}_L^{\vee}(a_R, a_R^{\dagger})$ by the following substitution,

$$
\hat{\mathcal{O}}_L(a^L,(a^L)^{\dagger}) = \sum_n c_{m,n}(a_k^L)^m(a_k^{L\dagger})^n \to \hat{\mathcal{O}}_L^{\vee}(a^R,a^{R\dagger}) \equiv \sum_n c_{m,n}(a_{-k}^R)^n(a_{-k}^{R\dagger})^m. \tag{4.29}
$$

Note that the coefficients $c_{m,n}$ are not converted to its complex conjugate. In particular, the field itself $\phi_L(t)$ is converted as

$$
\begin{aligned}
\phi_L(t) &= \int\frac{dk}{4\pi\omega_k}[a_k^L e^{i\omega_k t+ikr_*}+(a_k^L)^{\dagger}e^{-i\omega_k t-ikr_*}] &(4.30)\\
\to \phi_L^{\vee}(t) &= \int\frac{dk}{4\pi\omega_k}[(a_{-k}^R)^{\dagger}e^{i\omega_k t+ikr_*}+a_{-k}^R e^{-i\omega_k t-ikr_*}]. &(4.31)
\end{aligned}
$$

This has the same functional form as $\phi^R(t)$. We later interpret this field $\phi_L^{\vee}(t)$ as the (lower) Schwinger-Keldysh field and write as $\tilde{\phi}_R$ to distinguish the original right-wedge field ϕ_R. By this substitution, the expectation value can be written as

$$
\begin{aligned}
{}_K\langle 0|\hat{\mathcal{O}}_R\hat{\mathcal{O}}_L|0\rangle_K &= \frac{1}{Z}\mathrm{Tr}\left(e^{-\frac{\beta}{2}H_R}\ \hat{\mathcal{O}}_R\ e^{-\frac{\beta}{2}H_R}\ \hat{\mathcal{O}}_L^{\vee}\right) \\
&\equiv \langle\hat{\mathcal{O}}_R\hat{\mathcal{O}}_L^{\vee}\rangle_{\frac{\beta}{2},\frac{\beta}{2}}
\end{aligned}
\tag{4.32}
$$

In order to distinguish it from the ordinary finite temperature Green function, we have introduced the notation $\langle\cdots\rangle_{\frac{\beta}{2},\frac{\beta}{2}}$ as above.

If $\hat{\mathcal{O}}_L$ is made of a product of operators $\hat{\mathcal{O}}_L = \hat{A}_L\hat{B}_L$, it is converted as

$$
\hat{\mathcal{O}}_L = \hat{A}_L\ \hat{B}_L \to \hat{\mathcal{O}}_L^{\vee} = \hat{B}_L^{\vee}\ \hat{A}_L^{\vee}. \tag{4.33}
$$

Especially a care should be taken for the time evolution operator $U_L(t,t_0) \equiv T\exp\Big[-i\int_{t_0}^t dt\hat{H}_L(t)\Big]$. Following the above substitution rule, it is converted to

$$U_L(t,t_0) \equiv T\exp\left[-i\int_{t_0}^t dt\hat{H}_L(t)\right] \to U_L^\vee(t,t_0) = \tilde{T}\exp\left[-i\int_{t_0}^t dt\hat{H}_L^\vee(t)\right] \tag{4.34}$$

where $\tilde{T}$ is an anti-time ordering. For a hermitian Hamiltonian, $\hat{H}_L^\vee = \hat{H}_R$ is satisfied and

$$U_L^\vee(t,t_0) = \tilde{T}\exp\left[-i\int_{t_0}^t dt\hat{H}_R(t)\right] \tag{4.35}$$

Hence a Heisenberg operator $\hat{A}_L(x)$ is mapped to

$$\hat{A}_L(x) = U_L^\dagger(t_x,t_0)\hat{A}_L(t_0)U_L(t_x,t_0) \to \hat{A}_L^\vee(x) = U_L^\vee(t_x,t_0)\hat{A}_L^\vee(t_0)U_L^{\dagger\vee}(t_x,t_0). \tag{4.36}$$

The converted Heisenberg operator is evolved backward in time with the Hamiltonian $(-H_R)$.

From these considerations, an expectation value of a general operator including both of left and right operators can be represented as a path integral form of the right-handed fields;

$$\begin{aligned} {}_K\langle 0|\hat{\mathcal{O}}_R\hat{\mathcal{O}}_L|0\rangle_K &= \langle\hat{\mathcal{O}}_R\hat{\mathcal{O}}_L^\vee\rangle_{\frac{\beta}{2},\frac{\beta}{2}} \\ &= \int \mathcal{D}\phi_R\mathcal{D}\tilde{\phi}_R\ \mathcal{O}_R[\phi_R]\ \mathcal{O}_L^\vee[\tilde{\phi}_R]\exp\Big[iS[\phi_R]-iS[\tilde{\phi}_R]\Big]. \end{aligned} \tag{4.37}$$

Here ϕ_R represents the original right-wedge field while a new field $\tilde{\phi}_R(t)$ is introduced to represent the transformed operator $\mathcal{O}_L^\vee$. The minus sign in front of the action $S[\tilde{\phi}_R]$ comes from the backward time-evolution of $\mathcal{O}_L^\vee$. If we combine $\phi_R(t)$ and $\tilde{\phi}_R(t)$ together as a single $\phi_R(t)$ field along a doubled path depicted below, this expression is equivalent to the closed time path formalism of the real-time finite temperature field theory. The insertions of $\exp(-\beta H_R/2)$ can be represented as an evolution of time into the imaginary direction with $-\beta/2$ at both ends. Hence the path is given on the complex time plane as Fig. 3. The field on the lower line corresponds to the field in the left-wedge as $\phi_R(t-i\beta/2) = \phi_L(t)$.

An alternative interpretation is an analogy with the thermo field dynamics [23], another method to deal with the real-time finite temperature field theory. In this analogy, the operators in the left wedge can be regarded as the "tilde-fields" of thermo field dynamics [24].

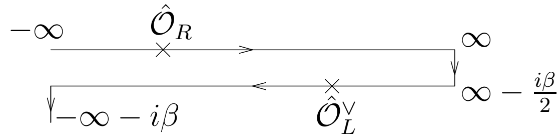

Figure 3: ϕ_R lives on the upper line while $\tilde{\phi}_R$ lives on the lower line. The time evolution is backward on the lower line.

4.3 Effective Equation of Motion in the Vicinity of the Horizon

Now we derive an effective equation of motion for the scalar field in the black hole background. Classically a dissipation term is induced since the energy is absorbed into the black hole horizon. In quantizing the system, a noise term will also be induced because of the Hawking radiation, and the system is effectively described by a Langevin equation.

The effect of the absorption can be described by imposing the ingoing boundary condition at the horizon $r = r_H$. Since, in the near horizon region, the system can be described by a set of 2-dim free fields satisfying $(\partial_t^2 - \partial_{r_*}^2)\phi_{(l,m)} = 0$, the ingoing boundary condition can be represented as

$$(\partial_t - \partial_{r_*})\phi_{(l,m)}(t, r = r_H) = 0. \tag{4.38}$$

The condition implies that there are no outgoing modes at the horizon, and violates the time reversal symmetry.

Since the scalar field is coupled to the gravitational field, if it is quantized, the chiral condition at the horizon seems to violate the general covariance by the quantum gravitational anomaly. The violation is compensated by the flux of the Hawking radiation [25, 26, 27]. In the following we will see that the quantization of the scalar field near the horizon naturally leads to the chiral condition (absorption) with the flux of Hawking radiation (noise term) at the horizon.

The method we will use is similar to the retarded-advanced (or Schwinger-Keldysh) formalism. The derivation of a Langevin equation is given by integrating fluctuating fields. (For a review, see, e.g. [28].)

4.3.1 Integrating Out the Environments

In obtaining the Langevin equation at the horizon, we need to integrate out certain kinds of environmental variables interacting with the *system* variable at the horizon. In order to do this, we first consider a stretched horizon at $r = r_H + \epsilon$ and treat the variables between the horizon ($r = r_H$) and the stretched horizon ($r = r_H + \epsilon$) as the environmental variables. Because of the quantum mixing of the wave functions in the left and right wedges (4.16),

the integration of these variables corresponds to an integration of the fields in the left wedge, which are classically hidden. In this way, we derive a Langevin equation at the stretched horizon. This equation is shown to be independent of the small parameter ϵ characterizing the position of the stretched horizon and we can take $\epsilon \to 0$ limit at last.

Since the Langevin equation we are going to derive is the equation of motion at the boundary of a rigion $[r_H, r_H + \epsilon]$, it is convenient to discretize the equation of motion near the horizon. In the tortoise coordinate r_* in which the equation of motion becomes free, the region is mapped to $[-\infty, r_*(r_H + \epsilon)]$. We divide the region into infinite segments as $(r_*)_n = r_*(r_H + \epsilon) + nd$ (for $n = 0, -1, -2, \cdots - \infty$) and set oscillators x_n on these lattice points. Here d is a lattice spacing in the tortoise coordinate. Discretized equations of motion for the scalar field are given by

$$\begin{cases} \ddot{x}_0 = -k(x_0 - x_1) + k(x_{-1} - x_0) - V_l((r_*)_0)x_0 \\ \vdots \\ \ddot{x}_{-n} = -k(x_{-n} - x_{-n-1}) + k(x_{-n+1} - x_{-n}) - V_l((r_*)_{-n})x_{-n}. \end{cases} \tag{4.39}$$

The continuum limit is given by taking $d \to 0$ limit with $kd^2 = 1$ and $\phi_{(l,m)}(t, (r_*)_n) = x_n(t)/\sqrt{d}$.

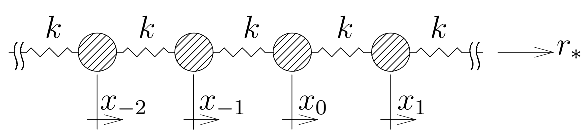

Figure 4: Discretized model of a scalar field in the near horizon region. The variable x_0 represents a variable at the stretched horizon.

Introducing the forward and the backward differentials in the tortoise coordinate r_*,

$$\Delta^+ x_n \equiv \frac{x_{n+1} - x_n}{d} \ , \ \Delta^- x_n \equiv \frac{x_n - x_{n-1}}{d}, \tag{4.40}$$

and using the relation,

$$-(x_n - x_{n+1}) + (x_{n-1} - x_n) = d\left(\Delta^+ - \Delta^-\right) x_n = d^2 \Delta^+ \Delta^- x_n, \tag{4.41}$$

we can write the discretized equation (4.39) for $n < 0$ as

$$\begin{aligned} \ddot{\phi}_{l,m}((r_*)_n) &= kd^2 \Delta^+ \Delta^- \phi_{l,m}((r_*)_n) - V_l((r_*)_n)\phi_{l,m}((r_*)_n) \\ \xrightarrow{d \to 0} \ \ddot{\phi}_{l,m}(r_*) &= \partial_{r_*}^2 \phi_{l,m}(r_*) - V_l(r_*)\phi_{l,m}(r_*). \end{aligned} \tag{4.42}$$

$V_l((r_*)_n)$ stands for the gravitational potential at $(r_*)_n$. It is proportional to $f(r) = 1 - 2M/r$ and vanishes in the near horizon region. Hence we neglect the potential term later in this section.

The normalization of the fields $\phi_{(l,m)}((r_*)_n) \equiv x_n/\sqrt{d}$ is determined from the action. With this normalization, the discretized action becomes the continuum one;

$$\begin{aligned} S &= -\int dt \left[\frac{1}{2} \sum_{n=-\infty}^{0} (\dot{x}_n)^2 - U(x_n) \right] \\ &= -\int dt \int_{-\infty}^{r_*(r_H+\epsilon)} dr_* \frac{1}{2} \left[(\dot{\phi}_{(l,m)}(r_*))^2 - (\partial_{r_*}\phi_{(l,m)}(r_*))^2 - V_l(r_*)(\phi_{(l,m)}(r_*))^2 \right]. \end{aligned} \tag{4.43}$$

Here the discretized potential U is defined by

$$U(x_n) \equiv \frac{1}{2} \sum_{n=-\infty}^{0} \left[k(x_{n+1} - x_n)^2 + V_l(r_n) x_n^2 \right]. \tag{4.44}$$

Note that, we define $d\sum_{n=-\infty}^{0} \to \int_{-\infty}^{r_*(r_H+\epsilon)} dr_*$ when $d \to 0$.

The full action is a sum of the fields in the left and the right wedges. As we saw in the previous section, the path integral containing both the left and right fields can be rewritten by a path integral of a right field on a closed time path. In the previous section, we have written the filed in the lower line by $\tilde{x}_R$. In the following, we use a unified notation and write x_R by x^1 and $\tilde{x}_R$ by x^2.

Previously we considered a path from $t = -\infty$ to ∞. It can be generalized to a path up to a finite time with fixed boundary conditions $x_0^I(t) = x_{\rm fin}^I$ $(I = 1, 2)$;

$$P[x_{\rm fin}^I, t] = \int^{x_0^I(t)=x_{\rm fin}^I} \prod_{n=-\infty}^{0} \mathcal{D}x_n^1 \mathcal{D}x_n^2 e^{iS[x_{-N}^1,\cdots,x_0^1,x_1^1] - iS[x_{-N}^2,\cdots,x_0^2,x_1^2]} \tag{4.45}$$

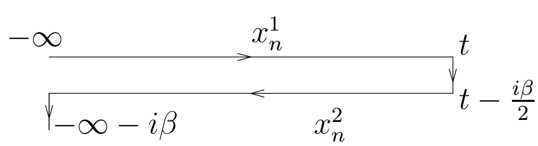

Figure 5: The values of the fields at the right ends of the paths, $x_n^I(t)$ $I = 1, 2$, are fixed in the path integral.

By integrating the environmental variables (fields between r_H and $r_H + \epsilon$), we have

$$P[x_{\rm fin}^I, t] = \int^{x_0^I(t)=x_{\rm fin}^I} \mathcal{D}x_0^1 \mathcal{D}x_0^2 e^{iS[x_0^1,x_1^1] - iS[x_0^2,x_1^2] + iS_{IF}[x_0^1,x_0^2]}. \tag{4.46}$$

The definition of the influence functional S_{IF} is schematically written by

$$e^{iS_{IF}[x_0^1,x_0^2]} \equiv \int \prod_{n=-\infty}^{-1} \mathcal{D}x_n^1 \mathcal{D}x_n^2 e^{iS[x_{-N}^1,\cdots,x_{-1}^1]-iS[x_{-N}^2,\cdots,x_{-1}^2]+iS_{int}[x_{-1}^1,x_0^1]-iS_{int}[x_{-1}^2,x_0^2]}, \tag{4.47}$$

where $S_{int}[x_{-1},x_0] = k/2\int dt(x_0 - x_{-1})^2$.

Since the *system* variables x_0^1, x_0^2 are coupled linearly with the environment variables, the influence functional $S_{IF}[x_0^1, x_0^2]$ has a Gaussian form

$$S_{IF}[x_0^1,x_0^2] = \frac{1}{2}\int dtdt' x_0^I(t) F_{IJ}(t,t') x_0^J(t'). \tag{4.48}$$

The Kernel function in the Schwinger-Keldysh formalism $F_{IJ}(t,t')$ can be obtained by taking derivatives of the influence functional as

$$\begin{aligned} F_{IJ}(t,t') &= \frac{1}{i}\frac{\delta^2}{\delta x_0^J(t')\delta x_0^I(t)} e^{iS_{IF}[x_0^1,x_0^2]}|_{x_0^I=0} \\ &= i(kd)^2 \begin{pmatrix} \langle T\Delta^+ x_{-1}^1(t)\Delta^+ x_{-1}^1(t')\rangle_{\frac{\beta}{2},\frac{\beta}{2}} & -\langle \Delta^+ x_{-1}^1(t)\Delta^+ x_{-1}^2(t')\rangle_{\frac{\beta}{2},\frac{\beta}{2}} \\ -\langle \Delta^+ x_{-1}^2(t)\Delta^+ x_{-1}^1(t')\rangle_{\frac{\beta}{2},\frac{\beta}{2}} & \langle \tilde{T}\Delta^+ x_{-1}^2(t)\Delta^+ x_{-1}^2(t')\rangle_{\frac{\beta}{2},\frac{\beta}{2}} \end{pmatrix}, \end{aligned} \tag{4.49}$$

The expectation means an integration over the environmental variables $x_{-\infty}^I, \cdots, x_{-1}^I$ $(I = 1,2)$. T stands for the time ordering, and $\tilde{T}$ is the anti-time ordering. As we saw in the previous subsection (4.37), these propagators are equal to the Schwinger-Keldysh (SK) ones with the path drawn in Fig.3. In the continuum limit, the discrete Green functions $d \times F_{IJ}(t,t')$ become the continuum counterpart

$$\begin{aligned} &F^{IJ}_{(l,m)(l',m')}(t,t') \\ &= i\partial_{r_*}\partial_{r'_*} \begin{pmatrix} \langle T\phi^1_{(l,m)}(t,r_*)\phi^1_{(l',m')}(t',r'_*)\rangle_{\frac{\beta}{2},\frac{\beta}{2}} & -\langle \phi^1_{(l,m)}(t,r_*)\phi^2_{(l',m')}(t',r'_*)\rangle_{\frac{\beta}{2},\frac{\beta}{2}} \\ -\langle \phi^2_{(l,m)}(t,r_*)\phi^1_{(l',m')}(t',r'_*)\rangle_{\frac{\beta}{2},\frac{\beta}{2}} & \langle \tilde{T}\phi^2_{(l,m)}(t,r_*)\phi^2_{(l',m')}(t',r'_*)\rangle_{\frac{\beta}{2},\frac{\beta}{2}} \end{pmatrix} |_{r=r'=r_H+\epsilon} \\ &\equiv \partial_{r_*}\partial_{r'_*} \begin{pmatrix} G^{11}_{(l,m)(l',m')}(t,r_*;t',r'_*) & G^{12}_{(l,m)(l',m')}(t,r_*;t',r'_*) \\ G^{21}_{(l,m)(l',m')}(t,r_*;t',r'_*) & G^{22}_{(l,m)(l',m')}(t,r_*;t',r'_*) \end{pmatrix} |_{r=r'=r_H+\epsilon} \end{aligned} \tag{4.50}$$

and the influence functional is given by

$$S_{IF}[\phi^1(r_H+\epsilon),\phi^2(r_H+\epsilon)] = \frac{1}{2}\int dtdt' \phi^I_{(l,m)}(t,r_H+\epsilon) F^{IJ}_{(l,m),(l',m')}(t,t') \phi^J_{(l',m')}(t',r_H+\epsilon). \tag{4.51}$$

Strictly speaking, the expectation in the Green functions should be evaluated at $(r_*)_{-1}$, but in the continuum limit it coincides with the position at the stretched horizon at $r_H+\epsilon$.

4.3.2 Vacuum Condition

From the previous discussions, we already knew that the Green functions in the Kruskal vacuum become identical with the Schwinger-Keldysh Green functions along the contour in Fig.3. We repeat the discussion for the case of the two point functions explicitly in the following. In order to calculate the influence functional, we need to specify the vacuum condition for the environmental variables, i.e. the Kruskal vacuum condition so that the physical quantities is regular in the Kruskal coordinates. We expand the scalar field by $u_k^{(1)}, u_k^{(2)}$ and its complex conjugates

$$\phi_{(l,m)}(t,r_*) = \int \frac{dk}{\sqrt{4\pi\omega_k}} \left[c^{(1)}_{k(l,m)} u_k^{(1)} + (c^{(1)}_{k(l,m)})^\dagger (u_k^{(1)})^* + c^{(2)}_{k(l,m)} u_k^{(2)} + (c^{(2)}_{k(l,m)})^\dagger (u_k^{(2)})^* \right], \tag{4.52}$$

with the canonical commutation relations

$$[c^{(1)}_{k(l,m)}, (c^{(1)}_{k'(l',m')})^\dagger] = (2\pi)2\omega_k \delta_{ll'}\delta_{mm'}\delta(k-k'), \tag{4.53}$$

$$[c_k^{(2)(l,m)}, (c_{k'}^{(2)(l',m')})^\dagger] = (2\pi)2\omega_k \delta_{ll'}\delta_{mm'}\delta(k-k'). \tag{4.54}$$

The correlators in the Kruskal vacuum become the following forms,

$$F^{AB}_{K,(l,m)(l',m')}(t,t') = \delta_{ll'}\delta_{mm'}\partial_{r_*}\partial_{r'_*} G_K^{AB}(t,r_*;t',r'_*)_{|r_*=r'_*} \tag{4.55}$$

where

$$G_K^{AB}(t,r_*;t',r'_*) = i \begin{pmatrix} \langle T\phi^R(t,r_*)\phi^R(t',r'_*)\rangle_K & \langle \phi^R(t,r_*)\phi^L(t',r'_*)\rangle_K \\ \langle \phi^L(t,r_*)\phi^R(t',r'_*)\rangle_K & \langle T\phi^L(t,r_*)\phi^L(t',r'_*)\rangle_K \end{pmatrix} \tag{4.56}$$

Here K means the expectation value in the Kruskal vacuum. As we saw, they are related to the Schwinger-Keldysh Green functions $F^{IJ}_{(l,m),(l',m')}(t,t') = \delta_{ll'}\delta_{mm'}\partial_{r_*}\partial_{r'_*}G^{IJ}(t,t')|_{r=r'=r_H+\epsilon}$ discussed in the previous section as

$$\begin{aligned} &\frac{1}{2}\int dtdt' \phi^A_{(l,m)}(t,r_H+\epsilon) F^{AB}_{K,(l,m),(l',m')}(t,t')\phi^B_{(l',m')}(t',r_H+\epsilon) \\ &= \frac{1}{2}\int dtdt' \phi^I_{(l,m)}(t,r_H+\epsilon) F^{IJ}_{(l,m),(l',m')}(t,t')\phi^I_{(l',m')}(t',r_H+\epsilon), \end{aligned} \tag{4.57}$$

where the Schwinger-Keldysh Green functions G^{IJ} are given by

$$\begin{aligned} G^{IJ}(t,r_*;t',r'_*) &= i \begin{pmatrix} \langle T\phi^1(t,r_*)\phi^1(t',r'_*)\rangle_{\frac{\beta}{2},\frac{\beta}{2}} & -\langle \phi^1(t,r_*)\phi^2(t',r'_*)\rangle_{\frac{\beta}{2},\frac{\beta}{2}} \\ -\langle \phi^2(t,r_*)\phi^1(t',r'_*)\rangle_{\frac{\beta}{2},\frac{\beta}{2}} & \langle \tilde{T}\phi^2(t,r_*)\phi^2(t',r'_*)\rangle_{\frac{\beta}{2},\frac{\beta}{2}} \end{pmatrix} \\ &= i\int \frac{dk}{4\pi\omega_k}\frac{1}{2\sinh(\pi\omega_k/\kappa)} \begin{pmatrix} M^{11}(t,t') & M^{12}(t,t') \\ M^{21}(t,t') & M^{22}(t,t') \end{pmatrix} e^{ik(r_*-r'_*)}. \end{aligned} \tag{4.58}$$

Non-diagonal entries have an extra minus sign with respect to eq.(4.56), since the ϕ^2 field is defined to evolve backward in time as in Fig.3.

Each component can be calculated as

$$\begin{aligned}
M^{11}(t,t') &= \theta(t-t')\left(e^{\frac{\pi\omega}{\kappa}}e^{-i\omega(t-t')}+e^{-\frac{\pi\omega}{\kappa}}e^{i\omega(t-t')}\right)\\
&\quad+\theta(t'-t)\left(e^{-\frac{\pi\omega}{\kappa}}e^{-i\omega(t-t')}+e^{\frac{\pi\omega}{\kappa}}e^{i\omega(t-t')}\right), \qquad (4.59)\\
M^{22}(t,t') &= \theta(t'-t)\left(e^{\frac{\pi\omega}{\kappa}}e^{-i\omega(t-t')}+e^{-\frac{\pi\omega}{\kappa}}e^{i\omega(t-t')}\right)\\
&\quad+\theta(t-t')\left(e^{-\frac{\pi\omega}{\kappa}}e^{-i\omega(t-t')}+e^{\frac{\pi\omega}{\kappa}}e^{i\omega(t-t')}\right),\\
M^{12}(t,t') &= M^{21}(t,t') = e^{-i\omega(t-t')}+e^{i\omega(t-t')}. \qquad (4.60)
\end{aligned}$$

It can be also rewritten in the following form,

$$G^{IJ}(t,r_*;t',r'_*) = \int\frac{dk_0dk}{(2\pi)^2}e^{-ik_0(t-t')+ik(r_*-r'_*)}$$
$$\begin{pmatrix} \frac{1}{-k_0^2+\omega_k^2-i\epsilon}+2\pi i n(\omega_k)\delta(-k_0^2+\omega_k^2) & -2\pi i\sqrt{n(1+n)}\delta(-k_0^2+\omega_k^2) \\ -2\pi i\sqrt{n(1+n)}\delta(-k_0^2+\omega_k^2) & \frac{-1}{-k_0^2+\omega_k^2+i\epsilon}+2\pi i n(\omega_k)\delta(-k_0^2+\omega_k^2) \end{pmatrix}, \qquad (4.61)$$

where $n(\omega_k) = 1/(e^{\beta_H\omega_k}-1)$, $\beta_H = 2\pi/\kappa$. The 2-2 component of the Green function coincides with the anti-time ordered finite temperature Green function, while the 1-1 component is the ordinary time ordered one.

In the conventional real-time finite-temperature field theory, the contour is usually taken as in the figure 6. The contour corresponds to considering an ordinary finite tem-

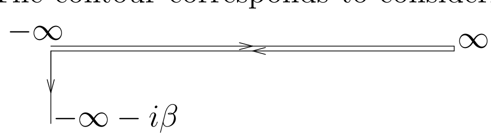

Figure 6: This path corresponds to the propagators which have non-diagonal entries (4.65).

perature Green function with the Boltzmann factor $e^{-\beta H}$ at the left-end;

$$\langle\hat{\mathcal{O}}^1\hat{\mathcal{O}}^2\rangle_\beta = \frac{1}{Z}\mathrm{Tr}\left(e^{-\beta H}\hat{\mathcal{O}}^1(t)\ \hat{\mathcal{O}}^2(t')\right) \qquad (4.62)$$

irrespective of whether these operators live on the upper or lower lines. On the other hand, the Green function (4.37) we are considering corresponds to taking a different contour as

drawn in Fig. 3. The fields on the lower line in these two contours are related by the similarity transformation

$$\phi^2_{\rm New}(t,x) = e^{-\beta H_R/2}\phi^2(t,x)e^{\beta H_R/2} = \phi^2(t+i\beta/2). \tag{4.63}$$

Here ϕ^2 and $\phi^2_{\rm New}$ are fields appearing in the formalism of Fig 3 and Fig 6 respectively. In the momentum representation, it is

$$\phi^2_{\rm New}(k) = e^{\frac{\beta k_0}{2}}\phi^2(k), \tag{4.64}$$

In terms of the new field, the Green function can be written in the following form,

$$\begin{aligned}
&G^{IJ}_{\rm New}(t,r_*;t',r'_*)\\
&= i\begin{pmatrix} \langle T\phi^1(t)\phi^1(t')\rangle_\beta & -\langle\phi^1(t)\phi^2_{\rm New}(t')\rangle_\beta \\ -\langle\phi^2_{\rm New}(t)\phi^1(t')\rangle_\beta & \langle\tilde{T}\phi^2_{\rm New}(t)\phi^2_{\rm New}(t')\rangle_\beta \end{pmatrix}\\
&= \int \frac{dk_0 dk}{(2\pi)^2} e^{-ik_0(t-t')+ik(r_*-r'_*)}\\
&\begin{pmatrix} \frac{1}{-k_0^2+\omega_k^2-i\epsilon} + 2\pi i n(\omega_k)\delta(-k_0^2+\omega_k^2) & -2\pi i\ \mathrm{sgn}(k_0)n(k_0)\delta(-k_0^2+\omega_k^2) \\ -2\pi i\ \mathrm{sgn}(k_0)(1+n(k_0))\delta(-k_0^2+\omega_k^2) & \frac{-1}{-k_0^2+\omega_k^2+i\epsilon} + 2\pi i n(\omega_k)\delta(-k_0^2+\omega_k^2) \end{pmatrix}.
\end{aligned} \tag{4.65}$$

4.3.3 Langevin equation at Stretched Horizon

The effective equation of motion at the stretched horizon can be obtained by taking a variation of the effective action $S[x_0^1] - S[x_0^2] + S_{IF}[x_0^1, x_0^2]$. In taking a continuum limit, a care should be taken since we have already integrated out the environmental field x_{-1}, and only the interaction with the outer variable x_1 appears in the effective action for x_0. The equation of motion for x_0^I becomes

$$\ddot{x}_0^I = -k(x_0^I - x_1^I) - \int dt' F^{IJ}(t,t')x_0^J(t'). \tag{4.66}$$

In the continuum limit ($d \to 0$) with kd^2 =fixed, the time derivative term drops and we have

$$\partial_{r_*}\phi^I_{(l,m)}(t) - \int dt' F^{IJ}_{(l,m)(l',m')}(t,t')\phi^J_{(l',m')}(t') = 0. \tag{4.67}$$

(Note that the discretized $d \times F^{IJ}$ becomes the continuum $F^{IJ}_{(l,m)(l',m')}$.) The dynamics seems to have disappeared in the effective equation at the stretched horizon, but we will see that another time derivative term (which is first order) is induced from the second term.

In order to show this, following the retarded-advanced formalism discussed below, we recombine the Schwinger-Keldysh fields, $\phi^1(x)$ and $\phi^2(x)$, into a *classical* variable $\phi^r(x)$ and a *fluctuating* variable $\phi^a(x)$. The interpretation of *classical* and *fluctuating* variables comes from an observation that the action $S[\phi^1] - S[\phi^2]$ has a dominant contribution in the path integral when the configuration of two fields coincide. As we saw in Fig. 3, the time axis of $\phi^1(t)$ differs from that of $\phi^2(t)$ by an amount of $\beta/2$ into the imaginary direction, and dominant configurations are given in terms of the redefined field (4.64) $\phi^2_{\text{New}}(t) = \phi^1(t + i\beta/2)$ in the following way. We define the *classical* and *fluctuating* fields as

$$\begin{cases} \phi^r_{(l,m)} = \frac{1}{\sqrt{2}}\left(\phi^1_{(l,m)} + \phi^2_{\text{New}(l,m)}\right) \\ \phi^a_{(l,m)} = \frac{1}{\sqrt{2}}\left(\phi^1_{(l,m)} - \phi^2_{\text{New}(l,m)}\right). \end{cases} \tag{4.68}$$

Propagators are transformed in this basis as

$$\left(\phi^1 \ \phi^2_{\text{New}}\right)\begin{pmatrix} G^{11}_{\text{New}} & G^{12}_{\text{New}} \\ G^{21}_{\text{New}} & G^{22}_{\text{New}} \end{pmatrix}\begin{pmatrix} \phi^1 \\ \phi^2_{\text{New}} \end{pmatrix} = (\phi^r \ \phi^a)\begin{pmatrix} 0 & G^A \\ G^R & 2iG^{\text{sym}} \end{pmatrix}\begin{pmatrix} \phi^r \\ \phi^a \end{pmatrix}. \tag{4.69}$$

where we have defined

$$G^R(t) = \frac{1}{2}(G^{11}_{\text{New}} - G^{12}_{\text{New}} + G^{21}_{\text{New}} - G^{22}_{\text{New}})(t) = i\theta(t)\langle[\phi(t), \phi(0)]\rangle \tag{4.70}$$

$$G^A(t) = \frac{1}{2}(G^{11}_{\text{New}} + G^{12}_{\text{New}} - G^{21}_{\text{New}} - G^{22}_{\text{New}})(t) = -i\theta(-t)\langle[\phi(t), \phi(0)]\rangle \tag{4.71}$$

$$G^{\text{sym}}(t) = -\frac{i}{4}(G^{11}_{\text{New}} - G^{12}_{\text{New}} - G^{21}_{\text{New}} + G^{22}_{\text{New}})(t) = \frac{1}{2}\langle\{\phi(t), \phi(0)\}\rangle \tag{4.72}$$

and used the relation $G^{11}_{\text{New}} + G^{12}_{\text{New}} + G^{21}_{\text{New}} + G^{22}_{\text{New}} = 0$. Because of this, the basis $(\phi^r \ \phi^a)$ are often called the retarded-advanced basis.

In terms of the r, a-fields the influence functional can be written as

$$\begin{aligned} S_{IF} = \int dt dt' \Big[& \phi^a_{(l,m)}(t)\partial_{r_*}\partial_{r'_*}G^R_{(l,m)(l',m')}(t,t')\phi^r_{(l',m')}(t') \\ & + i\phi^a_{(l,m)}(t)\partial_{r_*}\partial_{r'_*}G^{\text{sym}}_{(l,m)(l',m')}(t,t')\phi^a_{(l',m')}(t')\Big]. \end{aligned} \tag{4.73}$$

The derivative of the retarded Green function $\partial_{r_*}\partial_{r'_*}G^R$ satisfies

$$\partial_{r_*}\partial_{r'_*}G^R_{(l,m)(l',m')}(t,t')|_{r=r'=r_H+\epsilon} = -\delta_{ll'}\delta_{mm'}\partial_{t'}\delta(t-t'), \tag{4.74}$$

On the other hand, the symmetric Green function can be written as

$$G^{\text{sym}} = \int \frac{dk}{4\pi\omega_k}\left(n + \frac{1}{2}\right)\left(e^{-i\omega_k(t-t')} + e^{+i\omega_k(t-t')}\right)e^{ik(r_* - r'_*)}, \tag{4.75}$$

and its derivative becomes

$$\partial_{r_*}\partial_{r'_*}G^{\rm sym}_{(l,m)(l',m')}(t,t')|_{r=r'=r_H+\epsilon} = \delta_{ll'}\delta_{mm'}\int\frac{dk_0}{4\pi}\frac{k_0}{\tanh\frac{\beta k_0}{2}}e^{-ik_0(t-t')}$$
$$= \delta_{ll'}\delta_{mm'}\frac{1}{2\pi}\left[-\frac{\kappa^2}{4\sinh^2\frac{\kappa(t-t')}{2}}\right]. \quad (4.76)$$

for $t \neq t'$. The integral is divergent at $t = t'$. Since we are interested in the finite temperature effect, we regularize the symmetrized correlator by removing the κ-independent divergence (note $T = \kappa/2\pi$) as

$$K(t,t') \equiv \frac{1}{2\pi}\left[-\frac{\kappa^2}{4\sinh^2\frac{\kappa(t-t')}{2}} + \frac{1}{(t-t')^2}\right]. \quad (4.77)$$

Hence the action for the stretched horizon variable, which is a sum of S_{IF} and the interaction term with the neighboring variable x_1, becomes

$$S = \int dtdt'd\Omega r_\epsilon^2\Big[\phi^a(t,r_\epsilon,\Omega)\delta(t-t')(\partial_{t'}-\partial_{r^*})\phi^r(t',r_\epsilon,\Omega)$$
$$+ i\phi^a(t,r_\epsilon,\Omega)K(t,t')\phi^a(t',r_\epsilon,\Omega)\Big]. \quad (4.78)$$

Here, $r_\epsilon \equiv r_H + \epsilon$ appears with rewriting $\phi^r_{(l,m)}$ to ϕ^r. By integrating the fluctuating variable $\phi^a(t)$, (4.45) is written as

$$P[\phi^r_{\rm fin},t] =$$
$$\int^{\phi^r(t)=\phi^r_{\rm fin}}\mathcal{D}\phi^r\exp\left[-\frac{1}{4}\int dtdt'd\Omega r_\epsilon^2(\partial_t-\partial_{r^*})\phi^r(t,r_\epsilon,\Omega)K^{-1}(t,t')(\partial_{t'}-\partial_{r'^*})\phi^r(t',r_\epsilon,\Omega)\right] \quad (4.79)$$

It describes the effective dynamics at the stretched horizon. Note that the effective action contains a term which is odd under the time reversal transformation.

Instead of integrating out the fluctuating variable, we can introduce an auxiliary field $\xi(t)$ by

$$\exp\left(-\int dtdt'\phi^a_{(l,m)}(t)K(t,t')\phi^a_{(l,m)}(t')\right)$$
$$= \int\mathcal{D}\xi\exp\left(i\int dt\phi^a_{(l,m)}(t)\sqrt{2}\xi_{(l,m)}(t) - \frac{1}{2}\int dtdt'\xi_{(l,m)}(t)K^{-1}(t,t')\xi_{(l,m)}(t')\right). \quad (4.80)$$

Then the probability to see $\phi^r(t) = \phi^r_{\rm fin}$ at the stretched horizon is written in terms of the

scalar fields $\phi^{r,a}(t)$ and the auxiliary field ξ as

$$P[\phi^r_{\rm fin}, t] = \int^{\phi^r(t)=\phi^r_{\rm fin}} \mathcal{D}\phi^r \mathcal{D}\phi^a \mathcal{D}\xi \; e^{-\frac{1}{2}\int dtdt' \xi_{(l,m)}(t) K^{-1}(t,t') \xi_{(l',m')}(t')} e^{iS_{\rm eff}[\phi(t),\xi]}, \tag{4.81}$$

$$S_{\rm eff}[\phi(t),\xi] = \int dt \phi^a_{(l,m)}(t) \left[-\partial_{r_*}\phi^r_{(l,m)}(t) + \int dt' G^R_{(l,m)(l',m')}(t,t')\phi^r_{(l',m')}(t') + \sqrt{2}\xi_{(l,m)}(t) \right]. \tag{4.82}$$

The variation with respect to ϕ^a gives the equation of motion for ϕ^r

$$(\partial_t - \partial_{r_*})\phi^r_{(l,m)} + \sqrt{2}\xi_{(l,m)}(t) = 0, \tag{4.83}$$

with the Gaussian noise correlation

$$\langle \xi_{(l,m)}(t) \rangle = 0 \ , \ \langle \xi_{(l,m)}(t) \xi_{(l',m')}(t') \rangle = \delta_{ll'}\delta_{mm'} K(t,t'). \tag{4.84}$$

As expected, if we take the statistical average, the outgoing modes vanish in the averaged sense $\langle (\partial_t - \partial_{r_*})\phi^r \rangle = 0$, which means that there are only ingoing modes at the (stretched) horizon. The noise term can be considered as the effect of the Hawking radiation. In the next subsection, we compare the noise correlation obtained here with the flux of the Hawking radiation. The noise correlation is not white, and the memory effect remains with a time scale of the Hawking temperature $(t-t') \sim 1/\kappa = \hbar/2\pi T_H$. If we look at the dynamics of a time scale larger than it, we can approximate the noise as the following white noise

$$\langle \xi_{(l,m)}(t) \xi_{(l',m')}(t') \rangle \longrightarrow \delta_{ll'}\delta_{m,m'} \frac{\kappa}{2\pi} \delta(t-t') = \delta_{ll'}\delta_{m,m'} T_H \delta(t-t') \tag{4.85}$$

The above effective action is obtained previously based on the physical picture of the Hawking radiation [5] or a technique to reproduce the Schwinger-Keldysh formalism [7] in a setting of vibrating string in AdS black hole background. We have reproduced the same effective action by explicitly integrating the environmental variables between the horizon and the stretched horizon. Because of the mixing of the wave functions (4.16), the integration corresponds to an integration over the variables hidden in the horizon.

5 The Fluctuation Theorem for Black Holes and Matters

Now we apply the fluctuation theorem to the scalar field in the black hole background. Most generally, we must treat the whole system of the scalar field and the space-time

as a coupled quantum system, and backreactions to the space-time structure must be included. In our previous letter [29], we briefly sketched how to treat the metric degrees of freedom quantum mechanically in the path integral formalism and discussed the effect of the backreaction. In the present paper, in order to give a more systematic and complete investigation, we consider an easier situation, i.e. a scalar system in a fixed black hole background. We neglect effects of backreactions. Even if we adopt such a simplification, various interesting results follow the fluctuation theorem applied to our system, such as a proof of the generalized second law and a derivation of the Green-Kubo formula.

5.1 Discretized Equations outside the Stretched Horizon

The equation of motion of the scalar field $\phi^r(t, r_*)$ in the black hole background consists of the two coupled equations, namely, the effective equation at the stretched horizon $r = r_H + \epsilon$ and the bulk equation of motion outside the stretched horizon.

We put the scalar field in a box with a radius $r_B(> r_H)$ and impose a boundary condition at the outer boundary $\phi(t, r = r_B, \Omega) = 0$ in this subsection. Owing to the boundary condition, the scalar field is shown to be thermalized. Another merit of confining the system in a box is to stabilize the total system (even we take the backreaction into account [30] if the size of the box is not so large.) It thus justifies to choose an equilibrium distribution as an initial distribution for the matter field as we will do in the following. In a later section, we choose a different boundary condition to realize a steady state with a constant energy flux.

In order to apply the fluctuation theorem reviewed in section 3, we need to construct a Fokker-Planck equation which is local in time. In doing so, it is necessary to approximate the noise correlation (4.84) by the white noise (4.85). This approximation is valid for a longer time scale than $\hbar/2\pi T_H$. Though the validity is limited, we consider such an approximation in the present paper. The memory effect of the colored noise will be discussed later.

In the white noise approximation, the discretized equations are given by

$$\begin{cases} \gamma_0 \dot{x}_0 = -k(x_0 - x_1) - \sqrt{2}\xi_0 \ , \quad \langle \xi_0(t)\xi_0(t')\rangle = \gamma_0 T_H \delta(t-t') \\ \ddot{x}_1 = -k(x_1 - x_2) + k(x_0 - x_1) - V_l(r_1)x_1 \\ \vdots \\ \ddot{x}_N = -k(x_N - x_{N+1}) + k(x_{N-1} - x_N) - V_l(r_N)x_N \\ x_{N+1} \equiv 0. \end{cases} \tag{5.1}$$

The first line is the stochastic equation for the field at the inner boundary (stretched horizon) $r = r_\epsilon \equiv r_H + \epsilon$ with a noise term ξ. Note that the time derivative term originates

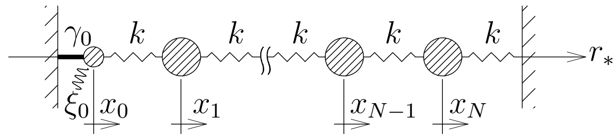

Figure 7: A schematic illustration of eq.(5.1). Each variables x_i are bound by spring with constant k. Only x_0 reserves a friction γ_0 and a noise ξ_0. The effects of $V_l(r_i)$ are not described on picture, it affects each variables as harmonic potential.

in the dissipation term induced by the interaction with the environmental variables. The last one is the boundary condition at the outer boundary $r = r_B$. The middle ones are the bulk equations of motion, and the field $\phi_{l,m}(r_*)$ between the stretched horizon and the outer boundary is discretized into N lattice points in the tortoise coordinate r_*. The continuum limit can be taken as before by taking $d \to 0$ (or $N \to \infty$) with the following conditions and replacements;

$$(N+1)d = r_*(r_B) - r_*(r_\epsilon),\ kd^2 \equiv 1\ ,\ \gamma_0 d \equiv 1,\ \phi^r_{(l,m)}((r_*)_i) \equiv x_i/\sqrt{d},\ \xi_{(l,m)} \equiv \sqrt{d}\xi_0, \tag{5.2}$$

where d is the lattice spacing in the tortoise coordinate r_*. The continuum equations in the bulk can be recovered as before and becomes

$$\ddot{\phi}^r_{(l,m)}(t, r_*) = \partial^2_{r_*}\phi^r_{(l,m)}(t, r_*) - V_l(r_*)\phi^r_{(l,m)}(t, r_*). \tag{5.3}$$

At the stretched horizon, the first equation of (5.1) can be written as

$$\begin{aligned} \gamma_0 \dot{x}_0 &= kd\Delta^+ x_0 - \sqrt{2}\xi_0 \\ \xrightarrow{d\to 0}\ \dot{\phi}^r_{(l,m)}(t, r_\epsilon) &= \partial_{r_*}\phi^r_{(l,m)}(t, r_\epsilon) - \sqrt{2}\xi_{(l,m)}, \end{aligned} \tag{5.4}$$

with noise correlation

$$\begin{aligned} \langle \xi_{(l,m)}(t)\xi_{(l',m')}(t')\rangle &= \delta_{ll'}\delta_{mm'} d\langle \xi_0(t)\xi_0(t')\rangle = \delta_{ll'}\delta_{mm'} d\gamma_0 T_H \delta(t-t') \\ &= \delta_{ll'}\delta_{mm'} T_H \delta(t-t'). \end{aligned} \tag{5.5}$$

5.2 Fluctuation Theorem for Scalar Field in Black Hole Background

From the set of the Langevin equations (5.1), we can construct the corresponding Fokker-Planck equation of $P(x_0, x_1, \cdots, x_N, v_1, \cdots, v_N, t)$ with $2N+1$ set of variables;

$$\partial_t P = \partial_{x_0}\left[\frac{1}{\gamma_0}\frac{\partial U}{\partial x_0}P + \frac{T_H}{\gamma_0}\partial_{x_0}P\right] + \sum_{i=1}^{N}\left[\partial_{x_i}(-v_i P) + \partial_{v_i}\left(\frac{\partial U}{\partial x_i}P\right)\right], \tag{5.6}$$

where we defined $U(x) \equiv \frac{1}{2}\sum_{i=1}^{N+1}\left[(\Delta^- x_i)^2 + V_l(r_i)x_i^2\right]$. In this expression, we introduced a redundant variable x_{N+1} for convenience, but eventually set $x_{N+1} = 0$. This equation has a solution describing an equilibrium distribution,

$$
\begin{aligned}
P^{\mathrm{eq}} &= Z^{-1} e^{-\frac{1}{T_H}\left[\frac{1}{2}\sum_{i=1}^{N} v_i^2 + U(x)\right]}, \\
Z &= \int d^{N+1}x d^N v e^{-\frac{1}{T_H}\left[\frac{1}{2}\sum_{i=1}^{N} v_i^2 + U(x)\right]}.
\end{aligned} \tag{5.7}
$$

A general solution to the Fokker-Planck equation can be formally represented in the path integral form as,

$$
P(x_0, x_i, \tau | x_0', x_i', t=0) \propto \int_{x'}^{x} \prod_{k=0}^{N} \mathcal{D}x_k e^{-\frac{1}{4\gamma_0 T_H}\int_{\Gamma_\tau} dt(\gamma_0 \dot{x}_0 - k(x_1 - x_0))^2} \prod_t \prod_{i=1}^{N} \delta(\ddot{x}_i + \tfrac{\partial U}{\partial x_i})|_{x_{N+1}\equiv 0}. \tag{5.8}
$$

The probability that a trajectory $\Gamma_\tau = \{(x_0', x_1', \cdots, x_N') \to (x_0, x_1, \cdots, x_N)\}$ is realized is given by

$$
P[\Gamma_\tau | x'] \propto e^{-\frac{1}{4\gamma_0 T_H}\int_{\Gamma_\tau} dt(\gamma_0 \dot{x}_0 - k(x_1 - x_0))^2} \prod_t \prod_{i=1}^{N} \delta(\ddot{x}_i + \tfrac{\partial U}{\partial x_i}). \tag{5.9}
$$

On the other hand, the probability that the reversed trajectory $\Gamma_\tau^* = \{(x_0, x_1, \cdots, x_N) \to (x_0', x_1', \cdots, x_N')\}$ is realized is given by

$$
P[\Gamma_\tau^* | x] \propto e^{-\frac{1}{4\gamma_0 T_H}\int_{\Gamma_\tau} dt(-\gamma_0 \dot{x}_0 - k(x_1 - x_0))^2} \prod_t \prod_{i=1}^{N} \delta(\ddot{x}_i + \tfrac{\partial U}{\partial x_i}). \tag{5.10}
$$

Hence the ratio of these two probabilities becomes

$$
\begin{aligned}
\frac{P_{(l,m)}[\Gamma_\tau | x']}{P_{(l,m)}[\Gamma_\tau^* | x]} &= \exp\left[\frac{1}{T_H}\int_{\Gamma_\tau} dt \dot{x}_0 k(x_1 - x_0)\right] \\
&= \exp\left[\frac{1}{T_H}\int_{\Gamma_\tau} dt \dot{\phi}^r_{(l,m)}(r_\epsilon)\partial_{r_*}\phi^r_{(l,m)}(r_\epsilon)\right].
\end{aligned} \tag{5.11}
$$

Here we have written the index for the angular momentum (l, m) explicitly. Summing over all the contributions from various partial waves (l, m), the ratio can be written as an integral over the stretched horizon;

$$
\begin{aligned}
\prod_{(l,m)} \frac{P_{(l,m)}[\Gamma_\tau | x']}{P_{(l,m)}[\Gamma_\tau^* | x]} &= \exp\left[\frac{1}{T_H}\int_{\Gamma_\tau} dt d\Omega\ r_\epsilon^2 \dot{\phi}^r(t, r_\epsilon, \Omega)\partial_{r_*}\phi^r(t, r_\epsilon, \Omega)\right] \\
&= \exp\left[\frac{1}{T_H}\int_{\Gamma_\tau} dt d\Omega\ r_\epsilon^2 T^r_t(t, r_\epsilon, \Omega)\right].
\end{aligned} \tag{5.12}
$$

Here we have used the definition of the energy-momentum tensor $T_t^r = \partial_t\phi^r\partial^r\phi^r = \partial_t\phi^r\partial_{r_*}\phi^r$. Logarithm of the ratio is proportional to the energy flux into the black hole $\Delta M[\Gamma_\tau] = \int_{\Gamma_\tau} dt d\Omega \ r_\epsilon^2 T_t^r(t, r_\epsilon, \Omega)$. Hence, by using the first law of black hole thermodynamics $T_H \Delta S_{BH}[\Gamma_\tau] = \Delta M[\Gamma_\tau]$, we can interpret this entropy production as an amount of difference of the black hole entropy during $t = 0 \sim \tau$,

$$\frac{P[\Gamma_\tau|x']}{P[\Gamma_\tau^*|x]} = \exp\left[\Delta S_{BH}[\Gamma_\tau]\right]. \tag{5.13}$$

In a more general setting, we can introduce an externally controlled parameter such as a variable mass term $m(t)$ in the potential $U(x; \lambda_t^F)$. Even in the presence of such an external parameter, the ratio can be shown to be given by the difference of the entropy,

$$\frac{P^F[\Gamma_\tau|x']}{P^R[\Gamma_\tau^*|x]} = \exp\left[\Delta S_{BH}[\Gamma_\tau]\right]. \tag{5.14}$$

In a case with time-dependent external parameters, the forward and the reversed protocols are generally different and we need to put F and R to distinguish them.

In order to apply the fluctuation theorem, we further multiply the above probabilities $P^F[\Gamma_\tau|x']$ (or $P^R[\Gamma_\tau^*|x]$) by probabilities for the initial distributions. As we discussed above we can assume that the system is in an equilibrium distribution at the external parameter λ_0^F (or λ_τ^F) with the Hawking temperature $P^{\rm eq}(x'; \lambda_0^F)$ (or $P^{\rm eq}(x; \lambda_\tau^F)$). Hence

$$\begin{aligned}
&\frac{P^F[\Gamma_\tau|x']P^{\rm eq}(x'; \lambda_0^F)}{P^R[\Gamma_\tau^*|x]P^{\rm eq}(x; \lambda_\tau^F)} \\
&= \exp\left[\Delta S_{BH}[\Gamma_\tau] - \beta\left(H[x'; \lambda_0^F] - H[x; \lambda_\tau^F]\right) + \beta\left(F(\lambda_0^F) - F(\lambda_\tau^F)\right)\right] \\
&= \exp\left[(\Delta S_{BH} + \Delta S_M)[\Gamma_\tau]\right].
\end{aligned} \tag{5.15}$$

Here, we defined the entropy difference of the matter by $\Delta S_M = -\beta(H[x'; \lambda_0^F] - H[x; \lambda_\tau^F]) + \beta(F(\lambda_0^F) - F(\lambda_\tau^F))$, where $H[x'; \lambda_0^F]$ is the total energy of the system at $t = 0$ with an external parameter λ_0^F and $F(\lambda_0^F)$ is the free energy defined by $Z(\lambda_0^F) = e^{-\beta F(\lambda_0^F)}$.

The fluctuation theorem is a direct consequence of the above key relation (5.15). As we saw in sec.3, it is straightforward to prove that

$$\frac{\rho^F(\Delta S_{BH} + \Delta S_M)}{\rho^R(-(\Delta S_{BH} + \Delta S_M))} = e^{\Delta S_{BH} + \Delta S_M}. \tag{5.16}$$

Here $\rho^F(\Delta S_{BH} + \Delta S_M)$ is the probability to observe a value of the total entropy production $\Delta S_{BH} + \Delta S_M$ with the forwardly controlled external parameter. The denominator is similarly defined as the probability to observe a negative value of the entropy production

in the reversed protocol. Since the right hand side is usually much bigger than 1, the numerator is generally much bigger than the denominator.

By integrating it, we have the Jarzynski equality;

$$\langle e^{-(\Delta S_{BH}+\Delta S_M)}\rangle = 1. \tag{5.17}$$

We observe that there must exist a path with $(\Delta S_{BH} + \Delta S_M) < 0$, i.e. an entropy decreasing path, otherwise the Jarzynski equality cannot be satisfied. As we saw in section 3, the generalized second law [31]

$$\langle(\Delta S_{BH} + \Delta S_M)\rangle \geq 0. \tag{5.18}$$

is derived using the Jensen inequality $\langle e^x\rangle \geq e^{\langle x\rangle}$. The above theorems (5.16) and (5.17) are also applicable to dynamical processes which are generally in non-equilibrium distributions, if the Fokker-Planck equation we have used is valid. As we noticed, the validity holds when the time scale of the dynamics is longer than the time scale of the inverse Hawking temperature $\hbar/(2\pi T_H)$. The condition is not always satisfied, and in such situations, we need to take effects of time-correlations of emissions.

5.3 Memory Effect and Quantum Corrections

In the previous sections, we have approximated the dynamics of the scalar fields by Langevin and Fokker-Planck equations. The approximation is valid when the noise correlation (4.84) can be replaced by the white noise and also the evolution of the scalar field ϕ^r is dominated by the classical path described by the Langevin equation. The first condition is violated for a shorter time scale than $\hbar/2\pi T_H$. The second condition is related to a justification of the Markovian process we have used. If we take $\hbar \to 0$ limit while keeping $T_H = \hbar\kappa/2\pi$ fixed, both conditions are satisfied. If these conditions are violated, we need to treat the system quantum mechanically without using the classical stochastic equations. More detailed analysis of such quantum corrections will be reported in a separate paper, and a brief sketch is given here.

We start from the action (4.78) at the stretched horizon. Before integrating out the variable ϕ^a, this gives an amplitude of the stretched horizon variables ϕ^1 and ϕ^2. But in terms of the variable ϕ^r, the path integral represents the evolution of a density matrix a la Schwinger-Keldysh, and the path integral (4.79) should be regarded as giving a probability, not an amplitude for the configuration ϕ^r. Based on this interpretation, we wrote it as P.

The classical limit with T_H fixed corresponds to replacing $K(t,t')$ by $T_H\delta(t-t')$. In this limit, the probability for a trajectory $\Gamma_\tau[\phi^r]$ with an initial value $\phi^r_{\rm ini}$ to be realized is given by

$$P[\Gamma_\tau|\phi^r_{\rm ini}] = \exp\left[-\frac{1}{4T_H}\int_{\Gamma_\tau} dtd\Omega r_\epsilon^2\left[(\partial_t-\partial_{r^*})\phi^r(t)\right]^2\right]\prod_{t,r>r_\epsilon,(l,m)}\delta\left[(\partial_t^2-\partial_{r_*}^2+V_l)\phi^r_{(l,m)}\right]. \tag{5.19}$$

The ratio of the forward and the backward probabilities is now given by

$$\frac{P[\Gamma_\tau|\phi^r_{\rm ini}]}{P[\Gamma^*_\tau|\phi^r_{\rm fin}]} = \exp\left[\frac{1}{T_H}\int_{\Gamma_\tau} dtd\Omega r_\epsilon^2\dot{\phi}^r(t)\partial_{r^*}\phi^r(t)\right] \tag{5.20}$$

and reproduces the previous result (5.12). The exponent is proportional to the energy flowing into the black hole across the horizon, and interpreted as an entropy increase of the black hole.

More generally, if we do not replace $K(t,t')$ by the white noise, the ratio becomes

$$\frac{P[\Gamma_\tau|\phi^r_{\rm ini}]}{P[\Gamma^*_\tau|\phi^r_{\rm fin}]} = \exp\left[\int dtdt'd\Omega r_\epsilon^2\dot{\phi}^r(t)K^{-1}(t,t')\partial_{r^*}\phi^r(t')\right], \tag{5.21}$$

which is nonlocal in time. By expanding the kernel in terms of derivatives of the delta functions, the exponent receives corrections to the energy flow. These corrections can be interpreted as flows of higher-spin currents (operators containing higher derivatives of fields) into the black hole. These terms vanish after taking a long-time average, but remain for a short time scale. Applying the fluctuation theorems with the nonlocal modification of the kernel, the entropy increase of the black hole receives higher derivative corrections. A geometric interpretation of these corrections is interesting.

Another important quantum correction is the violation of the Markovian assumption. If the path integral is not dominated by classical paths, we need to sum over all possible sequences of configurations at the level of amplitudes, instead of considering probabilities at the classical level. We also need to generalize the fluctuation theorem themselves at the fully quantum level. We hope to come back to these issues in near future.

5.4 Steady State Fluctuation Theorem in Black Holes

So far, we have applied the fluctuation theorem to a scalar field in an equilibrium distribution and disturbance around it. In realizing such a situation, we have put the black hole in a box with an adiabatic (insulating) wall. Instead we can consider a steady state with a constant (but very small) energy flow from a black hole to outside. This can be

realized by putting a black hole in a box in contact with a thermal bath with a slightly lower (or higher) temperature than the Hawking temperature. We set the temperature at the wall of the box as $T_w(< T_H)$. Since the energy flow is assumed to be very small, we neglect backreactions of the energy transfer to the black hole itself. The fluctuation theorem for the steady state is reviewed in the appendix C.

First we start from the following discretized form of the equations of motion;

$$\begin{cases} \gamma_0 \dot{x}_0 = -k(x_0 - x_1) - \sqrt{2}\xi_0 \ , \ \langle \xi_0(t)\xi_0(t')\rangle = \gamma_0 T_H \delta(t-t') \\ \ddot{x}_1 = -k(x_1 - x_2) + k(x_0 - x_1) - V_l(r_1)x_1 \\ \vdots \\ \ddot{x}_N = -\gamma\dot{x}_N - k(x_N - x_{N+1}) + k(x_{N-1} - x_N) - V_l(r_N)x_N - \sqrt{2}\xi \\ \qquad\qquad\qquad\qquad\qquad\qquad , \langle \xi(t)\xi(t')\rangle = \gamma T_w \delta(t-t') \\ x_{N+1} \equiv 0. \end{cases} \tag{5.22}$$

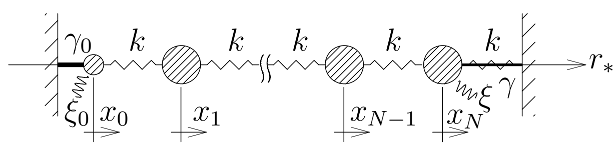

Figure 8: A discretized model of a scalar field in a box. The wall at $r = r_B$ is in contact with a thermal bath with temperature T_w slightly lower than the Hawking temperature. Then there is an energy flow from the black hole to the outer thermal bath.

The only difference from the previous model is the equation for the variable x_N at the wall. The variable in this model interacts with the thermal bath at temperature T_w. A redundant variable x_{N+1} is introduced for simplifying the above equation. The corresponding Fokker-Planck equation is given by

$$\begin{aligned} \partial_t P = \partial_{x_0}\left[\frac{1}{\gamma_0}\frac{\partial U}{\partial x_0}P + \frac{T_H}{\gamma_0}\partial_{x_0}P\right] + \sum_{i=1}^{N-1}\left[\partial_{x_i}(-v_i P) + \partial_{v_i}\left(\frac{\partial U}{\partial x_i}P\right)\right] \\ + \partial_{x_N}(-v_N P) + \partial_{v_N}\left(\gamma v_N + \frac{\partial U}{\partial x_N}P + \gamma T_w \partial_{v_N} P\right). \end{aligned} \tag{5.23}$$

The solution can be written in terms of the following path integral

$$\begin{aligned} P(x_{\rm fin}, \tau | x_{\rm ini}, 0) = \int_{x_{\rm ini}}^{x_{\rm fin}} \mathcal{D}\Gamma_\tau e^{-\frac{1}{4\gamma_0 T_0}\int_{\Gamma_\tau} dt\left(\gamma_0\dot{x}_0 + \partial_{x_0}U\right)^2} \prod_t \prod_{i=1}^{N-1} \delta\left(\ddot{x}_i + \partial_{x_i}U\right) \\ \times e^{-\frac{1}{4\gamma T_w}\int_{\Gamma_\tau} dt\left(\ddot{x}_N + \gamma\dot{x}_N + \partial_{x_N}U\right)^2}|_{x_{N+1}=0}. \end{aligned} \tag{5.24}$$

The ratio of the probabilities for a single partial wave with (l, m) is given by

$$\frac{P[\Gamma_\tau|x_i]}{P[\Gamma^*_\tau|x_f]} = \exp\left[-\frac{1}{T_H}\int_{\Gamma_\tau} dt\dot{x}_0\partial_{x_0}U - \frac{1}{T_w}\int_{\Gamma_\tau} dt\dot{x}_N(\ddot{x}_N + \partial_{x_N}U)\right]_{|x_{N+1}=0}. \tag{5.25}$$

In addition to the energy flow at the horizon, there is another contribution from the wall. The potential $U(x)$ is written as a sum of three terms;

$$U(x) = U_1(x_0, x_1, \cdots, x_{N-1}) + U_{12}(x_{N-1}, x_N) + U_2(x_N, x_{N+1}) \tag{5.26}$$

where

$$\begin{aligned}
U_1(x_0, x_1, \cdots, x_{N-1}) &= \frac{1}{2}\sum_{i=1}^{N-1}\left[\left(\frac{x_i - x_{i-1}}{d}\right)^2 + V_l(r_i)x_i^2\right], \\
U_{12}(x_{N-1}, x_N) &= \frac{1}{2}\left(\frac{x_N - x_{N-1}}{d}\right)^2, \\
U_2(x_N, x_{N+1}) &= \frac{1}{2}\left[\left(\frac{x_{N+1} - x_N}{d}\right)^2 + V_l(r_N)x_N^2\right].
\end{aligned} \tag{5.27}$$

We turn on the potential U_{12} at the wall during a time interval between $t = 0$ and $t = \tau$. This can be realized by introducing the external parameter controlling the potential U_{12} such as

$$U_{12}(x_{N-1}, x_N; \lambda^F_t) = \theta\left(\tfrac{\tau_-}{2} - |t - \tfrac{\tau}{2}|\right) U_{12}(x_{N-1}, x_N). \tag{5.28}$$

Then the variables $(x_0, x_1, \cdots, x_{N-1})$ are decoupled from x_N when $t < 0$ and $t > \tau$. Since the external thermal bath is decoupled for a long time during $t < 0$, the state can be considered in the equilibrium at $t = 0$. The ratio of the probabilities of the initial distributions is, hence, given by

$$\begin{aligned}
\frac{P^{\rm eq}(x_{\rm ini})}{P^{\rm eq}(x_{\rm fin})} = \exp\Bigg[&-\frac{1}{T_H}\left(\frac{1}{2}\sum_{i=1}^{N-1}(\dot{x}^2_{i,\rm ini} - \dot{x}^2_{i,\rm fin}) + U_1(x_{\rm ini}) - U_1(x_{\rm fin})\right) \\
&-\frac{1}{T_w}\left(\frac{1}{2}(\dot{x}^2_{N,\rm ini} - \dot{x}^2_{N,\rm fin}) + U_2(x_{\rm ini}) - U_2(x_{\rm fin})\right)\Bigg].
\end{aligned} \tag{5.29}$$

The second terms are canceled by the following terms in eq.(5.25)

$$\int_{\Gamma_\tau} dt\dot{x}_N(\ddot{x}_N + \partial_{x_N}U_2(x_N)) = \left[\frac{1}{2}m\dot{x}_N^2 + U_2(x_N)\right]^{\rm fin}_{\rm ini}. \tag{5.30}$$

The remaining terms in (5.29) is, of course, independent of the duration τ, and can be neglected compared to other terms in (5.25) that are proportional to τ.

As a result, if take the leading contributions in the large τ limit and neglect $\mathcal{O}(\tau^0)$ terms in the exponent, we obtain

$$\begin{aligned}
\frac{P[\Gamma_\tau|x_{\rm ini}]P^{\rm eq}(x_{\rm ini})}{P[\Gamma^*_\tau|x_{\rm fin}]P^{\rm eq}(x_{\rm fin})} &= \exp\left[-\frac{1}{T_H}\int_{\Gamma_\tau} dt\dot{x}_0\partial_{x_0}U_1 - \frac{1}{T_w}\int_{\Gamma_\tau} dt\dot{x}_N\partial_{x_N}U_{12}\right] \\
&= \exp\left[-\frac{1}{T_H}\int_{\Gamma_\tau} dt\dot{x}_0 kd\Delta^- x_0 - \frac{1}{T_w}\int_{\Gamma_\tau} dt\dot{x}_N kd\Delta^- x_N\right]. \qquad (5.31)
\end{aligned}$$

Because of the energy conservation for a steady state, we have the relation $\int dt\dot{x}_N\Delta^- x_N = -\int dt\dot{x}_0\Delta^- x_0$. In the contimuum limit $N \to \infty$ with the scalings explained before, the logarithm of the above ratio becomes

$$\begin{aligned}
&-\frac{1}{T_H}\int_{\Gamma_\tau} dt\dot{x}_0 kd\Delta^- x_0 - \frac{1}{T_w}\int_{\Gamma_\tau} dt\dot{x}_N kd\Delta^- x_N \\
&= (\beta_w - \beta_H)\int_{\Gamma_\tau} dt\partial_t\phi^r_{(l,m)}(t,r_\epsilon)\partial_{r_*}\phi^r_{(l,m)}(t,r_\epsilon) \equiv \Delta\beta\tau\bar{J}_{(l,m)}[\Gamma_\tau]. \qquad (5.32)
\end{aligned}$$

We have defined $\Delta\beta \equiv \beta_w - \beta_H$, which is positive from our assumption $T_H > T_w$. By summing all the contributions from the partial waves with (l,m), we have

$$\bar{J}[\Gamma_\tau] \equiv \frac{1}{\tau}\int_{\Gamma_\tau} dtd\Omega\ r_\epsilon^2\partial_t\phi^r(t,r_\epsilon,\Omega)\partial_{r_*}\phi^r(t,r_\epsilon,\Omega) = \frac{1}{\tau}\int_{\Gamma_\tau} dtd\Omega\ r_\epsilon^2 T^r_t(t,r_\epsilon,\Omega), \qquad (5.33)$$

where we have used the definition of the energy momentum tensor $T^r_t = \partial_t\phi^r\partial^r\phi^r = \partial_t\phi^r\partial_{r_*}\phi^r$. $\bar{J}[\Gamma_\tau]$ is a current flowing at the horizon out of the black hole. From the setting $T_H > T_w$, the averaged current is positive, but it can take both positive or negative values because of fluctuations of absorption and emission by the Hawking radiation.

We now have established the steady state fluctuation theorem in the black hole background as

$$\lim_{\tau\to\infty}\frac{1}{\tau}\left[\frac{\rho(\bar{J}_\tau,\Delta\beta)}{\rho(-\bar{J}_\tau,\Delta\beta)}\right] = \Delta\beta\bar{J}_\infty. \qquad (5.34)$$

For the definitions of ρ, see eq. C.6. The theorem can be restated in terms of a generating function $Z(\alpha_\tau,\Delta\beta)$, and leads to various relations between the response coefficients $L^{(1)}, L^{(2)}, \cdots$ defined by $\langle\bar{J}_\infty\rangle = L^{(1)}\Delta\beta + L^{(2)}/2(\Delta\beta)^2 + \cdots$ and correlator of currents $\langle J(t)J(t')\rangle$. For more details, see the appendix C.

In our case, these relations lead to the following relations;

$$L^{(1)} = \frac{1}{2}\int_0^\infty dt\int d\Omega\ r_\epsilon^4\langle T^r_t(t,r_\epsilon)T^r_t(0,r_\epsilon)\rangle_{|\Delta\beta=0} \qquad (5.35)$$

$$L^{(2)} = \lim_{\tau\to\infty}\frac{1}{2\tau}\int^\tau dtdt'\int d\Omega\ r_\epsilon^4\partial_{\Delta\beta}\langle T^r_t(t,r_\epsilon)T^r_t(t',r_\epsilon)\rangle_{|\Delta\beta=0} \qquad (5.36)$$

$$\vdots$$

The first relation is the Green-Kubo relation for the energy current flowing at the horizon $r = r_\epsilon$. The second one is a non-linear generalization, and the evaluation of the right hand side needs the derivative of the correlation function with respect to the temperature difference. This means that the information of the equilibrium distribution at $\Delta\beta = 0$ is not sufficient to obtain the non-linear response function of the current.

6 Summary

In this paper, we derived a stochastic equation with a dissipative term and a noise for a scalar field in a black hole background. The dissipation comes from the ingoing boundary condition at the horizon while the noise comes from the Hawking radiation. The stochastic equation can be derived by considering a stretched horizon and integrating variables between the horizon and the stretched horizon. The stochastic equation describes the dynamics of the scalar field in the limit $\hbar \to 0$ with the Hawking temperature $T_H = \hbar\kappa/2\pi$ kept finite. We then applied the non-equilibrium fluctuation theorems, developed in the statistical physics, to the above stochastic equation in the black hole background. We consider two cases. One is a scalar field confined in a box with an insulating wall. The system is relaxed to an equilibrium state at the Hawking temperature. The fluctuation theorem leads to the second law of the black hole thermodynamics after taking a thermal average. The other case is a scalar field in a box in contact with a heat bath with a different temperature from T_H. Then there is an energy flow between the horizon and the outer boundary. The fluctuation theorem leads to the Green-Kubo relation and its nonlinear generalizations.

We have used an approximation of replacing a nonlocal (colored) noise correlation by a white noise. We furthermore approximated the dynamical evolution of the scalar field by a classical Markovian process. These approximations are valid in the classical limit $\hbar \to 0$ with the Hawking temperature T_H fixed. In this sense, quantum effect is partially taken into account through the Hawking radiation. The results such as the ordinary second law of the black hole thermodynamics or the Green-Kubo relation are derived only in such approximations. As mentioned in Sec 5.3, the nonlocal noise correlation leads to a deviation of the black hole entropy appearing in the second law of thermodynamics. We hope to come back to these issues in near future.

Finally it is interesting to generalize our results for a scalar field to a vector or a gravitational field and obtain quantum corrections to the membrane action [4, 32]. The absorption of energy across the stretched horizon is known to give dissipative equations such as the Ohm's law or Navier-Stokes equation on the membrane. If the Hawking

radiation is included, these equations must receive quantum corrections as noise terms. Then the Hawking radiation may be interpreted as the anomaly inflow of the gravitational and gauge anomalies. In this sense, the quantum membrane action will be analogous to the edge state of quantum Hall effect. It is furthermore interesting if we can relate such a quantum membrane interpretation to the black hole entropy.

Acknowledgments

We would like to thank S. Zhang for collaboration at the early stage of the work. We thank the participants of the workshop "Towards New Developments in Field and String Theories" at RIKEN in December 17-19, 2010 for valuable discussions. We thank our colleagues, especially Y. Kitazawa and H. Kitamoto, for discussions, M. Hotta and M. Morikawa for inviting us to a wonderful workshop at Kusatsu in March 6-10, 2011. We furthermore acknowledge useful discussions with experts in the statistical physics, T. Sagawa, S. Sasa and H. Tasaki. The research by S.I. is supported in part by the Grant-in-Aid for Scientific Research (19540316) from MEXT, Japan. We are also supported in part by "the Center for the Promotion of Integrated Sciences (CPIS) " of Sokendai.

A Path integral form of the Fokker-Planck equation

In this appendix we derive the path integral form (2.11) of the solution to the Fokker-Planck equation;

$$\begin{aligned}\partial_t P(x,v,t|x_0,v_0,0) &= \hat{L}_{FT} P(x,v,t|x_0,v_0,0) \\ &= \partial_x\left(-vP\right) + \partial_v\left[\left(\frac{\gamma}{m}v + \frac{1}{m}\frac{\partial V}{\partial x}\right)P\right] + \partial_v^2\left(\frac{\gamma T}{m^2}P\right). \end{aligned} \tag{A.1}$$

For a small time-interval, it can be written as

$$\begin{aligned} &P(x,v,\Delta t|x_0,v_0,0) = e^{\Delta t \hat{L}_{FP}}\delta(x-x_0)\delta(v-v_0) \\ &\sim \int \frac{dk_x dk_v}{(2\pi)^2}\left[1+\Delta t\left[-v_0 ik_x + \left(\tfrac{\gamma}{m}v_0 + \tfrac{1}{m}\tfrac{\partial V(x_0)}{\partial x}\right)ik_v - \tfrac{\gamma T}{m^2}k_v^2\right]\right]e^{ik_x(x-x_0)+ik_v(v-v_0)} \\ &\sim \int \frac{dk_x dk_v}{(2\pi)^2}\exp\left[i\Delta t k_x\left(\tfrac{x-x_0}{\Delta t}-v_0\right) - \Delta t\tfrac{\gamma T}{m^2}\left(k_v - i\tfrac{m}{2\gamma T}\left(m\tfrac{v-v_0}{\Delta t}+\gamma v_0 + \tfrac{\partial V(x_0)}{\partial x}\right)\right)^2\right] \\ &\qquad\qquad \times \exp\left[-\tfrac{\Delta t}{4\gamma T}\left(m\tfrac{v-v_0}{\Delta t}+\gamma v_0+\tfrac{\partial V(x_0)}{\partial x}\right)^2\right] \\ &= \sqrt{\tfrac{2\pi m^2}{\Delta t \gamma T}}\delta(\dot{x}_0 - v_0)\exp\left[-\tfrac{\Delta t}{4\gamma T}\left(m\dot{v}_0 + \gamma v_0 + \tfrac{\partial V(x_0)}{\partial x}\right)^2\right] \end{aligned} \tag{A.2}$$

Then by using the Chapman-Kolmogorov equation $P(X_3|X_1) = \int dX_2 P(X_3|X_2)P(X_2|X_1)$ which is equivalent to an insertion of the complete set and integrating over v, we obtain the path integral form as follows;

$$P(x,t|x_0,0) = \int_{x(0)=x_0}^{x(t)=x} \mathcal{D}x \exp\left[-\tfrac{1}{4\gamma T}\int_0^t dt' \left(m\ddot{x} + \gamma\dot{x} + \tfrac{\partial V}{\partial x}\right)^2\right]. \tag{A.3}$$

If we use the Langevin equation (2.1), the path integral is equivalent to the noise average with the weight function in eq. (2.3).

B Noise correlation and Hawking radiation

The noise correlation induced in the effective equation of motion for the boundary field at the stretched horizon $r = r_H + \epsilon$ can be interpreted as the Hawking radiation. Here we first review the method to determine the energy-momentum tensor in the black hole background by using the trace anomaly of the energy-momentum tensor and the regularity condition at the horizon [33], and then generalize the method to determine higher spin currents [34, 35, 36, 37].

In two dimensions, the trace of the energy-momentum tensor of a single scalar field (i.e. the central charge is $c = 1$) has an anomaly term proportion to the scalar curvature R

$$T^\mu_\mu = \frac{1}{24\pi}R. \tag{B.1}$$

Writing the metric in the conformal gauge $ds^2 = e^{\varphi(u,v)}(-dudv)$, the equation becomes $T_{uv} = -\frac{1}{24\pi}\partial_u\partial_v\varphi$. By combining with the conservation of the energy-momentum tensor $\nabla_\mu T^\mu_\nu = 0$, derivatives of the EM tensor $\partial_v T_{uu}(u,v)$ and $\partial_u T_{vv}(u,v)$ can be written as follows;

$$\partial_v T_{uu} = \frac{1}{24\pi}\left[\partial_u^2\partial_v\varphi - (\partial_u\varphi)(\partial_u\partial_v\varphi)\right] \tag{B.2}$$

$$\partial_u T_{vv} = \frac{1}{24\pi}\left[\partial_v^2\partial_u\varphi - (\partial_v\varphi)(\partial_u\partial_v\varphi)\right]. \tag{B.3}$$

From these equations, we can define a (anti-) holomorphic quantity

$$t_{uu}(u) \equiv T_{uu} - \frac{1}{24\pi}\left[\partial_u^2\varphi - \frac{1}{2}(\partial_u\varphi)^2\right] \tag{B.4}$$

$$t_{vv}(v) \equiv T_{vv} - \frac{1}{24\pi}\left[\partial_v^2\varphi - \frac{1}{2}(\partial_v\varphi)^2\right]. \tag{B.5}$$

They are often called (anti-) holomorphic energy-momentum tensors, but their transformation laws are anomalous and not tensors in the exact sense. Actually, under a coordinate transformation from (u,v) to $(U=U(u), V=V(v))$, they transform as

$$t_{UU}(U) = \frac{1}{(\kappa U)^2}\left[t_{uu}(u) + \frac{1}{24\pi}\{U,u\}\right], \tag{B.6}$$

where $\{U,u\}$ is the Schwarzian derivative,

$$\{U,u\} \equiv \frac{\partial_u^3 U}{\partial_u U} - \frac{3}{2}\left(\frac{\partial_u^2 U}{\partial_u U}\right)^2. \tag{B.7}$$

In particularly, for the transformation from the Schwarzschild coordinates to the Kruskal ones, namely from (u,v) to $(U,V) = (-\kappa^{-1}e^{-\kappa u}, \kappa^{-1}e^{\kappa v})$, the Schwarzian derivative becomes $\{U,u\} = -\kappa^2/2$.

Now, we impose the regularity condition at the horizon. The energy momentum tensor T_{UU} must behave regularly near the future horizon $U=0$ in the regular coordinates, and so is $t_{UU}(U)$ since they are related regularly as (B.4). The regularity condition, hence, imposes that t_{uu} must behave as

$$t_{uu}(u\to\infty) = \frac{\kappa^2}{48\pi}. \tag{B.8}$$

If we neglect the effect of scatterings of the outgoing fluxes (namely in the absence of the gray body factor), we can extrapolate the above flux at the horizon to the outgoing flux at $r\to\infty$ as

$$T_{uu}(r\to\infty) = \frac{\kappa^2}{48\pi} = \frac{\pi}{12}T_H^2. \tag{B.9}$$

It is interpreted as the flux from the black body with the Hawking temperature T_H,

$$\int_0^\infty \frac{d\omega}{2\pi}\frac{\omega}{e^{\beta\omega}-1} = \frac{\pi}{12}T_H^2. \tag{B.10}$$

The transformation property of the holomorphic energy-momentum tensor can be also derived by considering the following point-splitting regularization,

$$\begin{aligned} :t_{uu}(u): &\equiv \lim_{\delta\to 0}\left[\partial_u\phi(u+\tfrac{\delta}{2})\partial_u\phi(u-\tfrac{\delta}{2}) - \langle\partial_u\phi(u+\tfrac{\delta}{2})\partial_u\phi(u-\tfrac{\delta}{2})\rangle\right] \\ &= \lim_{\delta\to 0}\left[\partial_u\phi(u+\tfrac{\delta}{2})\partial_u\phi(u-\tfrac{\delta}{2}) + \frac{1}{4\pi\delta^2}\right], \end{aligned} \tag{B.11}$$

where we have used the explicit form of the free boson propagator $\langle\phi(u)\phi(u')\rangle = -\ln(u-u')/4\pi$. From this definition, we can relate it to the energy momentum tensor regularlized

in the Kruskal (U) coordinate;

$$
\begin{aligned}
&: t_{uu}(u) : \\
&= \lim_{\delta\to 0}\left[\partial_u U(u+\tfrac{\delta}{2})\partial_u U(u-\tfrac{\delta}{2})\partial_U\phi(U(u+\tfrac{\delta}{2}))\partial_U\phi(U(u-\tfrac{\delta}{2})) + \frac{1}{4\pi\delta^2}\right] \\
&= \lim_{\delta\to 0}\left[\partial_u U(u+\tfrac{\delta}{2})\partial_u U(u-\tfrac{\delta}{2})\left(t_{UU}(U) - \frac{1}{4\pi}\frac{1}{(U(u+\frac{\delta}{2})-U(u-\frac{\delta}{2}))^2}\right) + \frac{1}{4\pi\delta^2}\right] \\
&= (\partial_u U)^2 : t_{UU}(U) :_K - \frac{1}{24\pi}\{U,u\}.
\end{aligned}
\tag{B.12}
$$

Namely, the Schwarzian derivative is nothing but the difference of the normal orderings in different coordinates.

The energy flux (which corresponds to the first moment of the thermal spectrum (B.10)) can be generalized to a flux of a higher spin current with a higher moment, and its generating function can be defined as a correlation function of the scalar field;

$$
\begin{aligned}
J(u,u+a) \equiv& \sum_{n=0}^{\infty}\frac{a^n}{n!} : \partial_u\phi(u)\partial^{n+1}\phi(u) : \\
&=: \partial_u\phi(u)\partial_u\phi(u+a) : \ .
\end{aligned}
\tag{B.13}
$$

The normal ordering : : is defined similarly to $t_{uu}(u)$ by

$$
: \partial_u\phi(u)\partial_u\phi(u) : \equiv \lim_{u'\to u}\left[\partial_u\phi(u)\partial_u\phi(u') - \langle\partial_u\phi(u)\partial_u\phi(u')\rangle\right]. \tag{B.14}
$$

Then we can show that $J(u,u+a)$ transforms under the coordinate transformation from u to $U(u)$ as

$$
J(u,u+a) = \partial_u U(u)\partial_u U(u+a) J(U(u),U(u+a)) + \frac{1}{4\pi}\left[-\frac{\kappa^2}{4\sinh^2\frac{\kappa a}{2}} + \frac{1}{a^2}\right]. \tag{B.15}
$$

Similarly to the energy flux discussed before, the regularity condition at the future horizon fixes the value of $J(u,u+a)$ at $U=0$ as

$$
J(u,u+a)|_{r=r_H} =: \partial_u\phi(u)\partial_u\phi(u+a) := \frac{1}{4\pi}\left[-\frac{\kappa^2}{4\sinh^2\frac{\kappa a}{2}} + \frac{1}{a^2}\right]. \tag{B.16}
$$

This can be interpreted as a correlation function of $\partial_u\phi(u)$ and $\partial_u\phi(u+a)$ on the Kruskal vacuum.

In Section 4.3.3, we have shown that the scalar field obeys a stochastic equation of motion

$$
\partial_u\phi(t-r^*)_{|r=r_H+\epsilon} = -\sqrt{2}\xi(t). \tag{B.17}
$$

at the stretched horizon. Since the equation is independent of the value of ϵ, we can safely take $\epsilon \to 0$ limit. Then the value of the generating function $J(u, u+a)$ for the higher spin fluxes discussed above is equivalent to the noise correlation $2\langle\xi(t)\xi(t+a)\rangle$ of the Langevin equation at the horizon. The functional forms are equal, though the coefficients are different by a factor 4. Reason for the discrepancy is now under study.

C The Steady State Fluctuation Theorem

In this appendix, we consider the fluctuation theorem for a steady state and derive the Green-Kubo formula.

Assume that we have two variable x_1, x_2, and each of them is in contact with a different thermal bath with temperature T_1 and T_2. We further assume that the dynamics is governed by the set of Langevin equations such as

$$
\begin{aligned}
m_1\dot{v}_1 + \frac{\partial V}{\partial x_1} + \gamma_1 v_1 = \xi_1 \quad &, \quad \langle \xi_1(t)\xi_1(t')\rangle = 2\gamma_1 T_1 \delta(t-t') \\
m_2\dot{v}_2 + \frac{\partial V}{\partial x_2} + \gamma_2 v_2 = \xi_2 \quad &, \quad \langle \xi_2(t)\xi_2(t')\rangle = 2\gamma_2 T_2 \delta(t-t').
\end{aligned}
\tag{C.1}
$$

Here, $V(x_1, x_2; \lambda_t^F)$ is an interaction potential between the two variables. The corresponding Fokker-Planck equation of the system can be obtained straightforwardly. The trajectory Γ_τ is also generalized as $\Gamma_\tau = \{x(t) = (x_1(t), x_2(t)) | x(0) = (x_1(0), x_2(0)) = (x_{\text{ini}}^1, x_{\text{ini}}^2)\}$. Then the solution to the Fokker-Planck equation gives probabilities of the forward and the reversed protocols, and the ratio is given by

$$
\frac{P^F[\Gamma_\tau|x_{\text{ini}}]}{P^R[\Gamma_\tau^*|x_{\text{fin}}]} = \exp\left[-\frac{1}{T_1}\int_{\Gamma_\tau} dt\dot{x}_1\left(m_1\ddot{x}_1 + \frac{\partial V(x;\lambda_t^F)}{\partial x_1}\right) - \frac{1}{T_2}\int_{\Gamma_\tau} dt\dot{x}_2\left(m_2\ddot{x}_2 + \frac{\partial V(x;\lambda_t^F)}{\partial x_2}\right)\right].
\tag{C.2}
$$

We have assumed that the two variables are decoupled before $t = 0$ and after $t = \tau$; the interaction potential V vanishes at $t < 0$ and $t > \tau$. The initial distribution of the total system is given by a product of the equilibrium distributions of each system $P^{\text{eq}}(x_{\text{ini}}) = P^{\text{eq}}(x_{\text{ini}}^1)P^{\text{eq}}(x_{\text{ini}}^2)$. The forward protocol is expressed as

$$
V(x;\lambda_t^F) = V_1(x_1) + V_2(x_2) + f(\lambda_t^F)V_{12}(x_1 - x_2)
\tag{C.3}
$$

where

$$
f(\lambda_t^F) = \theta\left(\frac{\tau_-}{2} - |\lambda_t^F - \frac{\tau}{2}|\right), \quad \lambda_t^F = t.
\tag{C.4}
$$

τ_- means $\tau - \epsilon$ for $0 < \epsilon \ll \tau$. Function $f(t)$ satisfies $f(t=0) = f(t=\tau) = 0$ and $f(0 < |t| < \tau) = 1$, so that the interaction switches on at $t = 0$ and off at $t = \tau$. This protocol has the reversal symmetry $f(\lambda_t^F) = f(\lambda_{\tau-t}^F)$.

In considering the large interval limit $\tau \to \infty$, the energy transfer such as $\int dt\dot{x}_1\partial_{x_1}V_{12}(x_1 - x_2)$ (or $\int dt\dot{x}_2\partial_{x_2}V_{12}(x_1-x_2)$) grows linearly in τ. On the other hand $\Delta E_1 = \int dt\dot{x}_1(m_1\ddot{x}_1 + \partial_{x_1}V_1(x_1)) = (\frac{1}{2}m_1\dot{x}_1^2 + V_1(x_1))_{t=\tau} - (\frac{1}{2}m_1\dot{x}_1^2+V_1(x_1))_{t=0}$ or $\Delta E_2 = \int dt\dot{x}_2(m_2\ddot{x}_2+\partial_{x_2}V_2(x_2))$ is at most $\mathcal{O}(\tau^0)$. If each system becomes stationary after taking $\tau \to \infty$, the change in the energy of each system vanishes. Hence we can drop the term $P^{\mathrm{eq}}(x_{\mathrm{ini}})/P^{\mathrm{eq}}(x_{\mathrm{fin}})$ and a contribution of ΔE_i in $P[\Gamma_\tau|x_{\mathrm{ini}}]/P[\Gamma_\tau^*|x_{\mathrm{fin}}]$ when we evaluate the quantity

$$\lim_{\tau\to\infty}\frac{1}{\tau}\ln\left(\frac{P[\Gamma_\tau|x_{\mathrm{ini}}]P^{\mathrm{eq}}(x_{\mathrm{ini}})}{P[\Gamma_\tau^*|x_{\mathrm{fin}}]P^{\mathrm{eq}}(x_{\mathrm{fin}})}\right). \tag{C.5}$$

In addition, we have $\int_{\Gamma_\tau} dt\dot{x}_1\partial_1 V_{12} \sim -\int_{\Gamma_\tau} dt\dot{x}_2\partial_2 V_{12} + \mathcal{O}(\tau^0)$. Therefore we can write the ratio of the probabilities only in terms of the energy current defined by $\bar{J}[\Gamma_\tau] \equiv \frac{1}{\tau}\int_{\Gamma_\tau} dt\dot{x}_1\partial_1 V_{12}$. Writing the temperature difference as $\Delta\beta \equiv \beta_2 - \beta_1$, we obtain the following relation;

$$\begin{aligned}
\rho(\bar{J}_\tau, \Delta\beta) &\equiv \int \mathcal{D}x P[\Gamma_\tau|x_{\mathrm{ini}}]P^{\mathrm{eq}}(x_{\mathrm{ini}})\delta(\bar{J}_\tau - \bar{J}[\Gamma_\tau]) \\
&\simeq \int \mathcal{D}x P[\Gamma_\tau^*|x_{\mathrm{fin}}]P^{\mathrm{eq}}(x_{\mathrm{fin}})e^{\tau\Delta\beta\bar{J}[\Gamma_\tau]}\delta(\bar{J}_\tau - \bar{J}[\Gamma_\tau]) \\
&= e^{\tau\Delta\beta\bar{J}_\tau}\int \mathcal{D}x P[\Gamma_\tau^*|x_{\mathrm{fin}}]P^{\mathrm{eq}}(x_{\mathrm{fin}})\delta(\bar{J}_\tau + \bar{J}[\Gamma_\tau^*]) \\
&= e^{\tau\Delta\beta\bar{J}_\tau}\rho(-\bar{J}_\tau, \Delta\beta).
\end{aligned} \tag{C.6}$$

The steady state fluctuation theorem can be written as

$$\lim_{\tau\to\infty}\frac{1}{\tau}\ln\left[\frac{\rho(\bar{J}_\tau, \Delta\beta)}{\rho(-\bar{J}_\tau, \Delta\beta)}\right] = \Delta\beta\bar{J}_\infty. \tag{C.7}$$

From this relation, we can derive the Green-Kubo relation and its non-linear generalizations. By using the generating function

$$Z(\alpha_\tau, \Delta\beta) \equiv \ln\left(\int_{-\infty}^{\infty} d\bar{J}_\tau e^{i\tau\bar{J}_\tau\alpha_\tau}\rho(\bar{J}_\tau, \Delta\beta)\right), \tag{C.8}$$

the steady state fluctuation theorem (C.6) can be recast into

$$Z(\alpha_\tau + i\Delta\beta, \Delta\beta) = Z(-\alpha_\tau, \Delta\beta). \tag{C.9}$$

Taking a derivative of both sides with respect to $\Delta\beta$ and setting $\Delta\beta = 0$, we have

$$\partial_{\Delta\beta}\left[Z(\alpha_\tau, 0) - Z(-\alpha_\tau, 0)\right] = -i\partial_{\alpha_\tau}Z(\alpha_\tau, 0). \tag{C.10}$$

The generating function can be expanded in terms of the correlators of $\bar{J}_\tau$ as

$$Z(\alpha_\tau, \Delta\beta) = \sum_{n=1}^{\infty} \frac{(i\tau\alpha_\tau)^n}{n!} G_n(\Delta\beta). \tag{C.11}$$

$G_n(\beta)$ gives a connected Green function of the averaged current

$$\bar{J}_\tau = \frac{1}{\tau}\int_0^\tau dt J(t). \tag{C.12}$$

Now the equation (C.10) is rewritten in the following form;

$$[1-(-1)^n]\,\partial_{\Delta\beta} G_n(0) = \tau G_{n+1}(0). \tag{C.13}$$

We further expand the one-point function of $\bar{J}_\tau$, which gives an expectation value of the current, with respect to the inverse temperature difference $\Delta\beta$ as

$$G_1(\Delta\beta) \equiv \sum_{m=0}^{\infty} \frac{L^{(m)}}{m!}(\Delta\beta)^m. \tag{C.14}$$

For $n = 0$, we have a trivial identity $G_1(0) = L^{(0)} = 0$. For $n = 1$, the Green-Kubo relation is derived;

$$\begin{aligned} L^{(1)} &= \frac{1}{2\tau}\int_0^\tau dtdt' \langle J(t)J(t')\rangle_{|\Delta\beta=0} \\ &\xrightarrow{\tau\to\infty} \frac{1}{2}\int_0^\infty dt \langle J(t)J(0)\rangle_{|\Delta\beta=0}. \end{aligned} \tag{C.15}$$

When $\Delta\beta = 0$, the system is described by the equilibrium distribution function $P^{\rm eq}(x) = e^{-\beta E_{\rm tot}(x)}/Z, \beta = \beta_1 = \beta_2$ and an expectation value of a function $F(x(t))$ is given by $\langle F(x(t))\rangle_{|\Delta\beta=0} = \int \mathcal{D}x P^{\rm eq}(x(t))F(x(t))$. In the large τ limit, the correlator $\langle J(t)J(t')\rangle_{|\Delta\beta=0}$ depends only on $(t-t')$.

We can also obtain the expression of $L^{(2)}, L^{(3)}, \cdots$ by taking further derivatives of the equation (C.9) with respect to $\Delta\beta$. For instance, we can derive

$$\begin{aligned} &\partial_{\Delta\beta}^2 [Z(\alpha_\tau,0) - Z(-\alpha_\tau,0)] = -i\partial_{\alpha_\tau}\partial_{\Delta\beta}[Z(\alpha_\tau,0)+Z(-\alpha_\tau,0)] \\ &\Rightarrow (1-(-1)^n)\,\partial_{\Delta\beta}^2 K_n(0) = \tau\left(1+(-1)^{n+1}\right)\partial_{\Delta\beta}K_{n+1}(0). \end{aligned} \tag{C.16}$$

For $n = 1$, we have

$$L^{(2)} = \lim_{\tau\to\infty}\frac{1}{2\tau}\int_0^\tau dtdt' \partial_{\Delta\beta}\langle J(t)J(t')\rangle_{|\Delta\beta=0}. \tag{C.17}$$

These nonlinear generalizations can be systematically obtained by using the steady state fluctuation theorems. We apply these expansion method to a system of a black hole and matter to obtain the Green-Kubo relation for a thermal current in the body of the paper.

References

[1] J. M. Bardeen, B. Carter, and S. Hawking, "The Four laws of black hole mechanics," *Commun.Math.Phys.* **31** (1973) 161–170.

[2] S. W. Hawking, "Particle Creation by Black Holes," *Commun. Math. Phys.* **43** (1975) 199–220.

[3] A. Strominger and C. Vafa, "Microscopic Origin of the Bekenstein-Hawking Entropy," *Phys. Lett.* **B379** (1996) 99–104, `arXiv:hep-th/9601029`.

[4] K. S. Thorne, R. H. Price, and D. A. MacDonald, "Black holes: The membrane paradigm," *Yale University Press* (1986) .

[5] J. de Boer, V. E. Hubeny, M. Rangamani, and M. Shigemori, "Brownian motion in AdS/CFT," *JHEP* **07** (2009) 094, `arXiv:0812.5112 [hep-th]`.

[6] A. N. Atmaja, J. de Boer, and M. Shigemori, "Holographic Brownian Motion and Time Scales in Strongly Coupled Plasmas," `arXiv:1002.2429 [hep-th]`.

[7] D. T. Son and D. Teaney, "Thermal Noise and Stochastic Strings in AdS/CFT," *JHEP* **07** (2009) 021, `arXiv:0901.2338 [hep-th]`.

[8] C. Herzog and D. Son, "Schwinger-Keldysh propagators from AdS/CFT correspondence," *JHEP* **0303** (2003) 046, `arXiv:hep-th/0212072 [hep-th]`.

[9] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, "Probability of second law violations in shearing steady states," *Phys. Rev. Lett.* **71** (1993) 2401–2404.

[10] C. Jarzynski, "Nonequilibrium Equality for Free Energy Differences," *Phys. Rev. Lett.* **78** (1997) 2690–2693, `arXiv:cond-mat/9610209`.

[11] R. P. Feynman and J. F. L. Vernon, "The theory of a general quantum system interacting with a linear dissipative system," *Ann. Phys.* **24** (1963) 118–173.

[12] A. O. Caldeira and A. J. Leggett, "Influence of dissipation on quantum tunneling in macroscopic systems," *Phys. Rev. Lett.* **46** (1981) 211.

[13] A. O. Caldeira and A. J. Leggett, "Quantum tunneling in a dissipative system," *Ann. Phys.* **149** (1983) 374–456.

[14] L.Onsager and S.Machlup, "Fluctuations and Irreversible Processes," *Phys. Rev.* **91** (1953) 1505–1512.

[15] D. J. Evans and D. J. Searles, "The Fluctuation Theorem ," *Adv. Phys.* **51** (2002) 1529 – 1585.

[16] T. Hatano and S. Sasa, "Steady-State Thermodynamics of Langevin Systems," *Phys. Rev. Lett.* **86** (2001) 3463–3466, `arXiv:cond-mat/0010405`.

[17] B. Piechocinska, "Information erasure," *Phys. Rev. A* **61** no. 6, (May, 2000) 062314.

[18] J. Kurchan, "A Quantum Fluctuation Theorem," (2000) , `arXiv:cond-mat/0007360 [cond-mat]`.

[19] G. Crooks, "Nonequilibrium measurement of free energy differences for microscopically reversible markovian systems," *J. Stat. Phys* **90** (1998) 1481–1487.

[20] S. W. Hawking and G. F. R. Ellis, "The large scale structure of space-time," *Cambridge University Press* (1975) .

[21] N. D. Birrell and P. C. W. Davies, "Quantum Fields in Curved Space," *Cambridge University Press* (1984) .

[22] W. G. Unruh, "Notes on black hole evaporation," *Phys. Rev.* **D14** (1976) 870.

[23] Y. Takahasi and H. Umezawa, "Thermo field dynamics," *Collect.Phenom.* **2** (1975) 55–80.

[24] W. Israel, "Thermo field dynamics of black holes," *Phys.Lett.* **A57** (1976) 107–110.

[25] S. P. Robinson and F. Wilczek, "A Relationship between Hawking radiation and gravitational anomalies," *Phys.Rev.Lett.* **95** (2005) 011303, `arXiv:gr-qc/0502074 [gr-qc]`.

[26] S. Iso, H. Umetsu, and F. Wilczek, "Hawking radiation from charged black holes via gauge and gravitational anomalies," *Phys.Rev.Lett.* **96** (2006) 151302, `arXiv:hep-th/0602146 [hep-th]`.

[27] S. Iso, H. Umetsu, and F. Wilczek, "Anomalies, Hawking radiations and regularity in rotating black holes," *Phys.Rev.* **D74** (2006) 044017, `arXiv:hep-th/0606018 [hep-th]`.

[28] E. A. Calzetta and B. L. Hu, "Nonequilibrium Quantum Field Theory," *Cambridge University Press* (2008) .

[29] S. Iso, S. Okazawa, and S. Zhang, "Non-Equilibrium Fluctuations of Black Hole Horizons," `arXiv:1008.1184 [gr-qc]`.

[30] G. W. Gibbons and M. J. Perry, "Black Holes and Thermal Green's Functions," *Proc. Roy. Soc. Lond.* **A358** (1978) 467–494.

[31] J. D. Bekenstein, "Generalized second law of thermodynamics in black hole physics," *Phys. Rev.* **D9** (1974) 3292–3300.

[32] M. Parikh and F. Wilczek, "An Action for black hole membranes," *Phys.Rev.* **D58** (1998) 064011, `arXiv:gr-qc/9712077 [gr-qc]`.

[33] S. M. Christensen and S. A. Fulling, "Trace anomalies and the hawking effect," *Phys. Rev. D* **15** no. 8, (Apr, 1977) 2088–2104.

[34] S. Iso, T. Morita, and H. Umetsu, "Higher-spin Currents and Thermal Flux from Hawking Radiation," *Phys.Rev.* **D75** (2007) 124004, `arXiv:hep-th/0701272 [hep-th]`.

[35] S. Iso, T. Morita, and H. Umetsu, "Fluxes of higher-spin currents and Hawking radiations from charged black holes," *Phys.Rev.* **D76** (2007) 064015, `arXiv:0705.3494 [hep-th]`.

[36] S. Iso, T. Morita, and H. Umetsu, "Higher-spin gauge and trace anomalies in two-dimensional backgrounds," *Nucl.Phys.* **B799** (2008) 60–79, `arXiv:0710.0453 [hep-th]`.

[37] S. Iso, T. Morita, and H. Umetsu, "Hawking radiation via higher-spin gauge anomalies," *Phys.Rev.* **D77** (2008) 045007, `arXiv:0710.0456 [hep-th]`.